

Access DB# _____

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

 Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: <u>Baderlye4998</u>	Type of Search	Vendors and cost where applicable
Searcher Phone #: _____	NA Sequence (#) _____	STN _____
Searcher Location: _____	AA Sequence (#) _____	Dialog _____
Date Searcher Picked Up: _____	Structure (#) _____	Questel/Orbit _____
Date Completed: <u>05-24-01</u>	Bibliographic _____	Dr. Link _____
Searcher Prep & Review Time: _____	Litigation _____	Lexis/Nexis _____
Clerical Prep Time: _____	Fulltext _____	Sequence Systems _____
Online Time: _____	Patent Family _____	WWW/Internet _____
	Other _____	Other (specify) <u>CGN</u>

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:42:44 ; Search time 63.83 Seconds
(without alignments)
111.049 Million cell updates/sec

Title: US-09-522-217-56_COPY_23_146

Perfect score: 659

Sequence: 1 QGPDRLRLRLHLDIVDEL.....KEFLERLKLWLLQKMIHQHLS 124

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0401.*

1:	/SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2:	/SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3:	/SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4:	/SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5:	/SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6:	/SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7:	/SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8:	/SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9:	/SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10:	/SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11:	/SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12:	/SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13:	/SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14:	/SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15:	/SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16:	/SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17:	/SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18:	/SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19:	/SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20:	/SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21:	/SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22:	/SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No.. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	659	100.0	146	21	B18624 A mouse zalphall 1
2	648	98.3	510	21	B18628 Amino acid sequenc
3	404	61.3	162	21	B18623 A human zalphall 1
4	394	59.8	519	21	B18627 Amino acid sequenc
5	152	23.1	32	21	B18626 Antigeninc peptide
6	126	19.1	40	21	B18625 ~Antigeninc peptide
7	80	12.1	1081	20	Y28482 cos-1 histidine ki
8	80	12.1	1081	20	W81600 Candida albicans C
9	74	11.2	670	21	B42416 Human ORF2180
10	73.5	11.2	632	17	R85300 Arabidopsis pathog
11	73.5	11.2	1269	17	W03659 RPP5 downy mildew

12	72.5	11.0	321	21	B44987 Human secreted pro
13	71.5	10.8	487	16	R71909 Human histamine H1
14	70.5	10.7	330	16	R79967 Fifth transmembran
15	70.5	10.7	487	16	R79965 Human histamine H1
16	70.5	10.7	874	19	W98698 H. pylori GHPO 686
17	69	10.5	651	21	Y82492 BYDV coat protein
18	69	10.5	724	21	Y82493 BYDV coat protein
19	68.5	10.4	114	16	R83309 Simian interleukin
20	68.5	10.4	114	17	W09101 Human mature epith
21	68.5	10.4	114	19	W39188 Human epithelium d
22	68.5	10.4	114	20	Y03759 Human epithelium d
23	68.5	10.4	114	21	Y52311 Mature human epith
24	68.5	10.4	122	17	R90842 Recombinant flag s
25	68.5	10.4	135	21	Y54825 Human interleukin-
26	68.5	10.4	162	16	R83436 Simian interleukin
27	68.5	10.4	162	16	R66926 Simian IL-15, Cer
28	68.5	10.4	162	17	W09100 Human epithelium d
29	68.5	10.4	162	17	W07254 Simian epithelium-
30	68.5	10.4	162	17	R98526 Simian interleukin
31	68.5	10.4	162	17	R92798 Mammalian interleu
32	68.5	10.4	162	19	W39187 Human epithelium d
33	68.5	10.4	162	20	Y03758 Human epithelium-d
34	68.5	10.4	162	21	Y78594 Simian interleukin
35	68.5	10.4	162	21	Y52310 Human epithelium-d
36	68.5	10.4	491	14	R36979 Histamine H1 recep
37	68	10.3	2749	12	R13887 Inositol-3-phospha
38	68	10.3	2749	21	Y77847 Mouse IP3 receptor
39	67	10.2	108	21	Y58205 Canine mature inte
40	67	10.2	132	21	Y58203 Canine interleukin
41	67	10.2	378	21	B43763 Human cancer assoc
42	67	10.2	406	19	W40035 Novel human protei
43	67	10.2	406	20	Y41731 Human PRO1075 prot
44	67	10.2	406	21	B44287 Human PRO1075 (UNQ
45	67	10.2	406	21	B24419 Human PRO1075 prot

ALIGNMENTS

RESULT	1
B18624	
ID	B18624 standard; Protein; 146 AA.
XX	
AC	B18624;
XX	
DT	22-JAN-2001 (first entry)
XX	
DE	A mouse zalphall ligand polypeptide.
XX	
KW	zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
KW	tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
XX	
OS	Mus musculus.
XX	
PN	WO200053761-A2.
XX	
PD	14-SEP-2000.
XX	
PF	09-MAR-2000; 2000WO-US06067.
XX	
PR	09-MAR-1999; 99US-0264908.
PR	11-MAR-1999; 99US-0265992.
PR	01-JUL-1999; 99US-0142013.
XX	
PA	(ZYNO) ZYMOGENETICS INC.
XX	
PI	Novak JE, Presnell SR, Sprecher CA, Foster DC, Gross JA;
PI	Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX	
DR	WPI; 2000-565600/52.
DR	N-PSDB; A75580.
XX	
PT	New human cytokine, designated zalphall ligand, useful for stimulating

PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX
 XX Disclosure; Page 222-223; 256pp; English.
 CC The present sequence represents a mouse zalphall ligand polypeptide,
 CC which is a cytokine. The zalphall ligand is useful for stimulating the
 CC proliferation and development of haematopoietic cells in vitro and in
 CC vivo. Zalphall ligand polynucleotides can be used as primers or probes
 CC for cloning the zalphall gene. The zalphall ligand is useful for
 CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
 CC used for treating leukaemias and lymphomas. Antagonists against zalphall
 CC ligand are useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 XX Sequence 146 AA;

Query Match 100.0%; Score 659; DB 21; Length 146;
 Best Local Similarity 100.0%; Pred. No. 2.4e-71;
 Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGPDRLRLRLHLDIVLVEQLKIYENDLDPDLLSAPQDVKGHCHEAFAACFQKAKLKPSNP 60
 DB 23 QGPDRLRLRLHLDIVLVEQLKIYENDLDPDLLSAPQDVKGHCHEAFAACFQKAKLKPSNP 82
 QY 61 GNNKTFIDLVLAOLRRRLPARRGGKKQKHIAKPCSDSYEKRTPKFLERLKWLQKMIH 120
 DB 83 gnnktfiidlvlaqlrrrlparrggkkqkhiakcpscdsyekrtpkfelerlkwllqkmih 142
 QY 121 QHLS 124
 DB 143 qhls 146

RESULT 2
 B18628
 ID B18628 standard; Protein; 510 AA.
 XX
 XX B18628;
 XX
 DT 22-JAN-2001 (first entry)
 XX
 DE Amino acid sequence of MBP-mouse zalphall ligand fusion in pTAP134.
 XX
 KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX
 OS Synthetic.
 OS Mus musculus.
 XX
 PN WO200053761-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06067.
 XX
 PR 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 XX (Zymo) ZYMOGENETICS INC.
 PA
 XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 XX

DR WPI; 2000-565600/52.
 DR N-PSDB; A75602.

XX New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX
 XX Example 31; Page 239-240; 256pp; English.

CC The present sequence represents a MBP-mouse zalphall ligand fusion in
 CC the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
 CC is useful for stimulating the proliferation and development of
 CC haematopoietic cells in vitro and in vivo. Zalphall ligand
 CC polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 XX Sequence 510 AA;

Query Match 98.3%; Score 648; DB 21; Length 510;
 Best Local Similarity 100.0%; Pred. No. 2.5e-69;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 PDRLRLRLHLDIVLVEQLKIYENDLDPDLLSAPQDVKGHCHEAFAACFQKAKLKPSNPGN 62
 DB 389 PDRLRLRLHLDIVLVEQLKIYENDLDPDLLSAPQDVKGHCHEAFAACFQKAKLKPSNPGN 448
 QY 63 NKTFIIDLVLAOLRRRLPARRGGKKQKHIAKPCSDSYEKRTPKFLERLKWLQKMIH 122
 DB 449 nktfiidlvlaqlrrrlparrggkkqkhiakcpscdsyekrtpkfelerlkwllqkmih 508
 QY 123 LS 124
 DB 509 ls 510

RESULT 3
 B18623
 ID B18623 standard; Protein; 162 AA.
 XX
 AC B18623;
 XX
 DT 22-JAN-2001 (first entry)
 XX
 DE A human zalphall ligand polypeptide.
 XX
 KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX
 OS Homo sapiens.
 XX
 PN WO200053761-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06067.
 XX
 PR 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 XX (Zymo) ZYMOGENETICS INC.

XX	Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI	Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX	
XX	WPI: 2000-565600/52.
DR	N-PSDB; A75552.
DR	
XX	
PT	New human cytokine, designated zalphall ligand, useful for stimulating
PT	the proliferation and/or development of haematopoietic cells in vitro
PT	and in vivo, and for treating tumourigenesis -
XX	
PS	Disclosure; Page 205-206; 256pp; English.
PS	
XX	
CC	The present sequence represents a human zalphall ligand polypeptide,
CC	which is a cytokine. The zalphall ligand is useful for stimulating the
CC	proliferation and development of haematopoietic cells in vitro and in
CC	vivo. Zalphall ligand polynucleotides can be used as primers or probes
CC	for cloning the zalphall gene. The zalphall ligand is useful for
CC	treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
CC	used for treating leukaemias and lymphomas. Antagonists against zalphall
CC	ligand are useful as research reagents for characterizing ligand-receptor
CC	interaction. Antagonists are also useful for inhibiting expansion,
CC	proliferation, activation and differentiation of cells involved in
CC	regulating hematopoiesis. The zalphall ligand may also be used to
CC	stimulate an immune response against B cell tumour, a virus, a parasite
CC	or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
CC	agonists and antibodies are also useful for the detection, diagnosis,
CC	prevention, and treatment of diseases associated with a zalphall ligand
CC	genetic defect.
XX	
XX	Sequence 162 AA: SQ

Query Match	61.3%;	Score 404;	DB 21;	Length 162;
Best Local Similarity	62.1%;	Pred. NO. 1.1e-40;		
Matches 77:	Conservative	19;	Mismatches 28;	Indels 0;
	Gaps	0;		

[illegible]

RESULT	4	
B18627		
ID	B18627	standard; Protein; 519 AA.
XX		
AC	B18627;	
XX		
DT	22-JAN-2001	(first entry)
XX		
DE	Amino acid sequence of MBP-human zalphall ligand fusion in pTAP126.	
XX		
KW	zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;	
KW	tumourigenesis; leukaemia; hematopoiesis; B cell tumour.	
XX		
OS	Synthetic.	
OS	Homo sapiens.	
XX		
PN	WO200053761-A2.	
XX		
PD	14-SEP-2000.	
XX		
PF	09-MAR-2000; 2000WO-US060657.	
XX		
PR	09-MAR-1999; 99US-0264908.	

11-MAR-1999; 99US-0265992.
01-JUL-1999; 99US-0142013.
(ZYMO) ZYMOGENETICS INC.
Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
Johnston JV, Nelson AU, Dillon SR, Hammond AK;
WPI; 2000-565600/52.
N-PSDB; A75599.
New human cytokine, designated zalphall1 ligand, useful for stimulating
the proliferation and/or development of haematopoietic cells in vitro
and in vivo, and for treating tumourigenesis -
Example 31; Page 233-235; 256pp; English.
The present sequence represents a MFP-human zalphall1 ligand fusion in
the plasmid pTAP126. Zalphall1 ligand is a cytokine. The zalphall1 ligand
is useful for stimulating the proliferation and development of
haematopoietic cells in vitro and in vivo. Zalphall1 ligand
polynucleotides can be used as primers or probes for cloning the
zalphall1 gene. The zalphall1 ligand is useful for treating tumourigenesis.
A zalphall1 ligand-saporin fusion toxin may be used for treating
leukaemias and lymphomas. Antagonists against zalphall1 ligand are
useful as research reagents for characterizing ligand-receptor
interaction. Antagonists are also useful for inhibiting expansion,
proliferation, activation and differentiation of cells involved in
regulating hematopoiesis. The zalphall1 ligand may also be used to
stimulate an immune response against B cell tumour, a virus, a parasite
or a bacterium. The zalphall1 polypeptides, polynucleotides, antagonists,
agonists and antibodies are also useful for the detection, diagnosis,
prevention, and treatment of diseases associated with a zalphall1 ligand
genetic defect.
Sequence 519 AA:

Query Match	59.8%	Score 394;	DB 21;	Length 519;
Best Local Similarity	62.0%	Pred. No. 8, 1e-39;		
Matches 75; Conservative	19;	Mismatches 27;	Indels	Gaps 0;

Qy	4	DRLIRLHLLDIVEQULKIYENDLDPELLSAPQDVKGHCHEHAAPACFQKAKLPSPNGNN	63
		: : : : : : : : : : : : :	
Db	390	drhmrmrqldivdqlknyvndlvpeflpapedvetncwsafscfqaklsantggn	449
		: : : : : : : : : : : :	
Qy	64	KTFIDLVAQLRRLPLPARGGKKOKHIAKPCSDSYBKRTPKFLERLKWLQKMIHQHL	123
		: : : : : : : : : : : : : : : : : : :	
Db	450	eriinvsikkjkrkppstnagrrqkhrktcpsdcsyekppkeflerfkallqmkmhql	509
		: : : : : : : : : : : :	
Qy	124	S 124	
Db	510	s 510	
		.	

RESULT	5	
BI8626		
ID	BI8626 standard; Peptide; 32 AA.	
XX		
XX	AC	
XX	BI8626;	
XX		
DT	22-JAN-2001 (first entry)	
XX		
XX	Antigenic peptide derived from a human zalphall ligand polypeptide.	
DE		
XX		
KW	zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;	
KW	tumourigenesis; leukaemia; hematopoiesis; B cell tumour.	
XX		
XX	Homo sapiens.	
OS		
XX		
PN	WC200053761-A2.	
XX		
PD	14-SEP-2000.	


```

XX PF 09-MAR-2000; 2000WO-US06067.
XX XX
XX PR 09-MAR-1999; 99US-0264908.
XX PR 11-MAR-1999; 99US-0265992.
XX PR 01-JUL-1999; 99US-0142013.
XX PA (ZYMO ) ZYMOGENETICS INC.
XX XX
XX PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Gross JA;
XX PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX DR WPI; 2000-565600/52.
XX XX
XX PT New human cytokine, designated zalphall ligand, useful for stimulating
XX PT the proliferation and/or development of haematopoietic cells in vitro
XX PT and in vivo, and for treating tumourigenesis -
XX XX
XX PS Example 34; Page 227; 256pp; English.
XX XX
XX CC The present sequence was used to raise antibodies, and is derived from
XX CC a human zalphall ligand polypeptide. Zalphall ligand is a cytokine.
XX CC The zalphall ligand is useful for stimulating the proliferation and
XX CC development of haematopoietic cells in vitro and in vivo. Zalphall
XX CC ligand polynucleotides can be used as primers or probes for cloning the
XX CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
XX CC A zalphall ligand-saporin fusion toxin may be used for treating
XX CC leukemias and lymphomas. Antagonists against zalphall ligand are
XX CC useful as research reagents for characterizing ligand-receptor
XX CC interaction. Antagonists are also useful for inhibiting expansion,
XX CC proliferation, activation and differentiation of cells involved in
XX CC regulating hematopoiesis. The zalphall ligand may also be used to
XX CC stimulate an immune response against B cell tumour, a virus, a parasite
XX CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
XX CC agonists and antibodies are also useful for the detection, diagnosis,
XX CC prevention, and treatment of diseases associated with a zalphall ligand
XX CC genetic defect.
XX SQ Sequence 32 AA;

Query Match 23.1%; Score 152; DB 21; Length 32;
Best Local Similarity 87.5%; Pred. No. 2.5e-11;
Matches 28; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 93 CPSCDSYEKRPKEFLERLKLQKMIHOHLS 124
Db 1 cpscdsyekkpkfelfkllqkmihnlis 32

RESULT 6
B18625
ID B18625 standard; Peptide; 40 AA.
XX AC B18625;
XX XX
XX DT 22-JAN-2001 (first entry)
XX DE
XX DE Antigenic peptide derived from a human zalphall ligand polypeptide.
XX KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
XX KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
XX OS Homo sapiens.
XX XX
XX PN WO200053761-A2.
XX PD
XX PD 14-SEP-2000.
XX XX
XX PF 09-MAR-2000; 2000WO-US06067.
XX XX
XX PR 09-MAR-1999; 99US-0264908.
XX PR 11-MAR-1999; 99US-0265992.

PR 01-JUL-1999; 99US-0142013.
XX PA (ZYMO ) ZYMOGENETICS INC.
XX XX
XX PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Gross JA;
XX PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX DR WPI; 2000-565600/52.
XX XX
XX PT New human cytokine, designated zalphall ligand, useful for stimulating
XX PT the proliferation and/or development of haematopoietic cells in vitro
XX PT and in vivo, and for treating tumourigenesis -
XX XX
XX PS Example 34; Page 227; 256pp; English.
XX XX
XX CC The present sequence was used to raise antibodies, and is derived from
XX CC a human zalphall ligand polypeptide. Zalphall ligand is a cytokine.
XX CC The zalphall ligand is useful for stimulating the proliferation and
XX CC development of haematopoietic cells in vitro and in vivo. Zalphall
XX CC ligand polynucleotides can be used as primers or probes for cloning the
XX CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
XX CC A zalphall ligand-saporin fusion toxin may be used for treating
XX CC leukemias and lymphomas. Antagonists against zalphall ligand are
XX CC useful as research reagents for characterizing ligand-receptor
XX CC interaction. Antagonists are also useful for inhibiting expansion,
XX CC proliferation, activation and differentiation of cells involved in
XX CC regulating hematopoiesis. The zalphall ligand may also be used to
XX CC stimulate an immune response against B cell tumour, a virus, a parasite
XX CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
XX CC agonists and antibodies are also useful for the detection, diagnosis,
XX CC prevention, and treatment of diseases associated with a zalphall ligand
XX CC genetic defect.
XX SQ Sequence 40 AA;

Query Match 19.1%; Score 126; DB 21; Length 40;
Best Local Similarity 64.1%; Pred. No. 4.5e-08;
Matches 25; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 4 DRLLRLRLHLDIVLQKLYIYENDLDPHLSAPQDVKGHC 42
Db 2 drhmirmrqlidivdqlknyvndlvpeflpapedvetnc 40

RESULT 7
Y28482
ID Y28482 standard; Protein; 1081 AA.
XX AC Y28482;
XX XX
XX DT 12-OCT-1999 (first entry)
XX DE
XX DE cos-1 histidine kinase.
XX XX
XX KW cos-1; osmosensing histidine kinase; fungal cell wall;
XX KW antifungal compound.
XX OS Candida albicans.
XX XX
XX PN US5939306-A.
XX PD
XX PD 17-AUG-1999.
XX PF 16-APR-1997; 97US-0843530.
XX PR 16-APR-1997; 97US-0843530.
XX XX
XX PA (CALY ) CALIFORNIA INST OF TECHNOLOGY.
XX PA (UYTE-) UNIV TECHNOLOGY CORP.
XX XX
XX PI Agnan J, Alex LA, Selitrennikoff C, Simon MI;
XX XX

```


XX PN W0200058473-A2.
XX PD 05-OCT-2000.
XX PF 31-MAR-2000; 2000WO-US08621.
XX PR 31-MAR-1999; 99US-0127607.
XX PR 02-APR-1999; 99US-0127636.
XX PR 05-APR-1999; 99US-0127728.
XX PR 30-MAR-2000; 2000US-0540763.
XX PA (CURA-) CURAGEN CORP.
XX PI Shimkets RA, Leach M;
XX PR WPI; 2000-602362/57.
XX DR N-PSDB; C76625.
XX PT Novel nucleic acids and peptides derived from open reading frame X,
XX PT useful for treating e.g. cancers, proliferative disorders,
XX PT neurodegenerative disorders and cardiovascular disease -
XX PS Claim 11; Page 3556-3558; 5507pp; English.
XX CC C74446 to C77606 encode the proteins given in B40237 to B43397, which
XX CC represent the human ORFX open reading frames 1 to 3161. The ORFX
XX CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
XX CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
XX CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
XX CC cardiac; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
XX CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
XX CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
XX CC sequences can be used for determining the presence of or predisposition
XX CC to, or preventing or treating pathological conditions associated with an
XX CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
XX CC proteins in gene therapy vectors. The proteins and nucleic acids may be
XX CC used to treat cancers, proliferative disorders, neurodegenerative
XX CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
XX CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
XX CC storage, systemic lupus erythematosus, severe combined immunodeficiency
XX CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
XX CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
XX CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
XX CC enhance coagulation; to inhibit thrombosis; and as a contraceptive.
XX SQ Sequence 670 AA;

Query Match 11.2%; Score 74; DB 21; Length 670;
Best Local Similarity 27.3%; Pred. No. 3;
Matches 35; Conservative 14; Mismatches 45; Indels 34; Gaps 6;

QY 6 LLIRLRHLIDIVEQLKIYENDLDP-----LLSAPQDVKGHCHEAFAFCQAKLPSN 59
Db 184 lfthrhgvcvagskppifvevdpgrcftiklsarkdhegsdcyprvc-----pnn 236
QY 60 PCNNKTFIIDVAQLRRLLPARRGGKKQKHIACKP-----SC-----DSVEKRTPKFELE 109
Db 237 pscppllrmleahl-----kecehi-kcphsksgyctfgnqdtyletcrfe 285
QY 110 RLKWLQK 117
Db 286 glkeflqq 293

RESULT 10
R85300
ID R85300 standard; Protein; 632 AA.
XX AC R85300;
XX DT 14-APR-1996 (first entry)

XX DE Arabidopsis pathogen resistance gene RPP5 protein.
XX KW Pathogen resistant; RPP5; tomato; C. fulvum; Avr 4; Avr 9; fungal;
XX KW leaf mould; variegation.
XX OS Arabidopsis sp.
XX PN W09531564-A2.
XX PD 23-NOV-1995.
XX PF 11-MAY-1995; 95WO-GB01075.
XX PR 07-APR-1995; 95GB-0007232.
XX PR 11-MAY-1994; 94GB-0009394.
XX PR 23-DEC-1994; 94WO-GB02812.
XX PR 31-MAR-1995; 95GB-0006638.
XX PA (GATS-) GATSBY CHARITABLE FOUND.
XX PI Hammond-Kosack KE, Jones DA, Jones JDG;
XX DR WPI; 1996-010949/01.
XX DR N-PSDB; T06308.
XX PT Increasing plant pathogen resistance by induction of variegation
XX PT may lead to acquired resistance to a broad range of pathogens.
XX PS Disclosure; Page 89-90; 131pp; English.
XX CC R85300 is the arabidopsis pathogen resistance gene RPP5 protein. In a
XX CC new method the RPP5 gene is expressed highly in genetic constructs which
XX CC may be used to impart a broad range of pathogen resistance by induction
XX CC of variegation, to transgenic plants (or parts or propagules of plants)
XX CC causing such constructs. RPP5 imparts resistance to the disease
XX CC caused by the leaf mould fungal pathogen Cladosporium fulvum.
XX CC C. fulvum contains avirulence (Avr) genes that confer recognition by
XX CC defence mechanisms to attack the disease.
XX CC N.B. The amino acid sequence given here and that given as SEQ ID 10 in
XX CC the specification are different, see features table and T06308.
XX SQ Sequence 632 AA;

Query Match 11.2%; Score 73.5; DB 17; Length 632;
Best Local Similarity 24.1%; Pred. No. 3.2;
Matches 33; Conservative 23; Mismatches 42; Indels 39; Gaps 6;

QY 4 DRLIRLRHLIDIVEQLKIYENDLDP-----SAPQDVKGHCHEAFAFCQAKL 55
Db 326 drqlk-aheidvyeqlpsqglalkmisyafigkdsppddfke-----lafevael 377
QY 56 KPSNP-----GNKKTFIIDVAQLRR-----RLPARRGGKKQKHIACKP 94
Db 378 vgsplglsvlgsslkgrdkewkmmprlndsdskkleetlrvgydrlnkknelfkci 437
QY 95 SC--DSVEKRTPKFELE 109
Db 438 acfngfkvsnvkelle 454

RESULT 11
W03659
ID W03659 standard; Protein; 1269 AA.
XX AC W03659;
XX DT 19-FEB-1997 (first entry)
XX DE RPP5 downy mildew resistance protein.
XX DT

KW	Arabidopsis; RPP5; disease-resistance; downy mildew;
KW	Peronospora parasitica; leucine-rich repeat; transgenic plant;
KX	crop improvement.
OS	Arabidopsis landrace Landsberg erecta.
XX	
XK	Key Location/Qualifiers
FH	Peptide 104..109
FT	/note= "Peptide W03660 (claim 22) used to construct primer T37477 (claim 23)"
FT	437..441
FT	Peptide /note= "Peptide W06331 (claim 22), used to construct primer T37478 (claim 23)"
FT	
XX	
PN	WO9631608-A1.
PN	
PD	10-OCT-1996.
XX	
PF	09-APR-1996; 96WO-GB00849.
PR	07-APR-1995; 95GB-0007232.
PA	(INNE-) INNES CENT INNOVATIONS LTD JOHN.
PI	Coleman M, Daniels MJ, Jones JDG, Parker J, Szabo V;
XX	WPI; 1996-465029/46.
DR	N-PSDB; T37476.
XX	
PT	Isolated Arabidopsis pathogen resistance gene RPP5 - for prodn. of transgenic plants esp. resistant to downy mildew fungus
PT	Claim 1; Fig 2; 59pp; English.
XX	
CC	This sequence represents an Arabidopsis RPP5 protein, which confers disease-resistance against the downy mildew fungus (peronospora parasitica). The sequence includes leucine-rich repeat regions, characteristic of many pathogen-resistance genes. The sequence shows strong homology to pathogen-resistance proteins N (from tobacco, conferring rust resistance) and L6 (from flax, conferring rust resistance), including regions involved in nucleotide binding (kinase-Ia (P-loop), kinase-2 and Kinase-3a domains). Primers T37477-78, corresponding to a conserved region between peptides W03660-61, may be used to identify other resistance genes in plants. The RPP5 gene may be expressed in a transgenic plant to confer disease-resistance.
XX	
SQ	Sequence 1269 AA;
Query Match	11.2%; Score 73.5; DB 17; Length 1269;
Best Local Similarity	24.1%; Pred. No. 7.8;
Matches	33; Conservative 23; Mismatches 42; Indels 39; Gaps 6;
OY	4 DRLLRLRHLDIVQLKIYNELDPELL-----SAPPDKHGCEHAFAFCPOAKL 55 :: :: :: :: :: :: :: :: :: :: :: 326 drqlk-aheidlvyeqvkpsglaikmsiyafkgdspddfe-----lafevael 377
Dz	56 KPSNP-----GNKKTFIDLVLAQLRR-----RLPARRGKGKQHHTAKCP 94 :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: 378 vsglpiglsvglsslgkrdkdwvkmprlrdsdckieetirvgdydrlnknreifklc1 437
OY	95 SC--DSYEKRTPKEFLF 109 : : :: 438 acfngfkysnvkelte 454
RESULT 12	
B44987	ID B44987 standard; Protein; 321 AA.
XX	
AC	B44987;
XX	

DT	12-FEB-2001 (first entry)
XX	
DE	Human secreted protein encoded by gene 19 homologue.
XX	
XK	Secreted protein; human; immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial; virucide; fungicide; ophthalmological; gene therapy; treatment; autoimmune disease; infection; cardiovascular disorder; nervous system disorder; ocular disorder; wound healing; epithelial cell proliferation; skin aging; mental state; transplantation; metabolism modulation.
XX	Homo sapiens.
OS	WO200055200-A1.
PN	21-SEP-2000.
XX	
PD	09-MAR-2000; 2000MO-USO6042.
PF	12-MAR-1999; 99US-0124143.
PR	03-DEC-1999; 99US-0168663.
XX	(HUMA-) HUMAN GENOME SCI INC.
PA	Rosen CA, Ruben SM, Komatsoulis G;
PI	WPI; 2000-656008/63.
XX	
DR	Isolated human secretory proteins, nucleic acids encoding them and antibodies directed against them, useful for diagnosing and treating disorders related to the proteins such as cancer, Alzheimer's disease and Parkinsons -
XX	Disclosure; Page 422-423; 453pp; English.
PS	This invention describes a novel isolated polypeptide (I) and its encoding nucleic acid molecule (II) which have immnosupressive, antiarthritic, antirheumatic, antiproliferative, cytosstatic, cardiant, vasotrophic, cerebroprotective, nootropic, neuroprotective, antibacterial, virucide, fungicide and ophthalmological activity and which can be used for gene therapy. (I) and (II) are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. (I) and (II) are also used in diagnosing a pathological condition or susceptibility to a pathological condition. The antibodies to (II) can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioluminoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides are used to modulate mammalian metabolism, to change mammals' mental state or physical state by influencing biorhythms circadian rhythms, depression tendency for violence tolerance for pain, reproductive capabilities, hormonal or endocrine levels, appetite, libido, memory, stress or other cognitive qualities, as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrates, vitamins, minerals, cofactors or other nutritional components.
XX	
SQ	Sequence 321 AA;
Query Match	11.0%; Score 72.5; DB 21; Length 321;
Best Local Similarity	26.8%; Pred. No. 1.8;
Matches	26; Conservative 12; Mismatches 32; Indels 27; Gaps 27;

QY 35 PQDVKGH-----CEHAFACFQAKLK---PSNPGNKKFTIIDLVAAQ----- 73
 Db 211 pselkxhmthtgekpyqcycifrcadqsnlkthikxhgnlpykcehpcpafgdere 270
 QY 74 LRRRLPARRGGKKQKHIACPCSDSYEKRTPKFELER 110
 Db 271 lqrhldlfqghkth-----qcphecd--hkstnsddlkr 301

RESULT 13

R71909
 ID R71909 standard; Protein; 487 AA.
 XX
 AC R71909;
 XX
 DT 29-NOV-1995 (first entry)
 XX
 DE Human histamine H1 receptor.
 XX
 KW Human histamine H1 receptor; polymorphism detection; treatments;
 KW drug design.
 XX
 OS Homo sapiens.
 XX
 PN JP07067654-A.
 XX
 PD 14-MAR-1995.

XX 03-SEP-1993; 93JP-0219544.
 XX
 PR 03-SEP-1993; 93JP-0219544.
 XX

PA (MIRP) MITSUBISHI YUKA BCL KK.
 XX
 DR WPI; 1995-143848/19.
 DR N-PSDB; Q87915.

XX Human histamine H1 receptor gene and protein - used in a method
 PT for the detection of histamine polymorphism
 PT
 PS Claim 1; Pages 9-12; 13pp; Japanese.
 XX

CC Q87915 encodes R71909 the human histamine H1 receptor. The DNA
 CC fragment can be used in a new method for the detection of
 CC histamine polymorphisms, it may also be used in the development
 CC of drugs for the treatment of histamine associated conditions.
 XX

SQ Sequence 487 AA;

Query Match 10.8%; Score 71.5; DB 16; Length 487;
 Best Local Similarity 28.3%; Pred. No. 4;
 Matches 26; Conservative 13; Mismatches 26; Indels 27; Gaps 5;

QY 21 KIYENDLPDLISAPQDVGHCHSE-----AFAFCQKAKLKPSNP-GNNKFTIIDLVAAQL 74
 Db 212 klykg-----vrqhcqhrelnrlslpsfseikrlpenkpgdakkpkespwev 259

QY 75 RRLPARRGGKKQKHIACPCSDSYEKRTPK 106
 Db 260 lkrkpkdagggs---vlksp-----qtpke 282

RESULT 14

R79967
 ID R79967 standard; Protein; 330 AA.
 XX

AC R79967;
 XX

DT 11-JAN-1996 (first entry)
 XX

DE Fifth transmembrane region of histamine H1 receptor.

XX

KW Histamine H1 receptor; human; bovine; transformation; drug screening;
 KW pNIV3604B; calcium ion mobilisation; inositol 1,4,5-triphosphate;
 KW CHO cell; guanine nucleotide regulatory protein; G protein;
 KW hybridisation.

OS Homo sapiens.
 XX

PN GB2283239-A.
 XX

PD 03-MAY-1995.
 XX

PF 29-OCT-1993; 93GB-00223353.
 XX

PR 29-OCT-1993; 93GB-00223353.
 XX

PA (UNIO) UCB SA.
 XX

PI Bollen A, Gillard M, Moguilevsky N, Noyer M, Varsalona F;
 XX

DR WPI; 1995-157301/21.
 DR N-PSDB; T04235.

XX New nucleic acid encoding human histamine H1 receptor - useful
 PT diagnostically and for screening receptor binding drugs
 PT
 XX Example 1; Page 26-29; 49pp; English.

XX This sequence represents the fifth transmembrane region of histamine H1
 CC receptor. This sequence was isolated after screening a lambda-tagged human
 CC lung cDNA library with the sequence shown in T04234. The human histamine
 CC H1 receptor cDNA sequence was constructed using the sequence
 CC encoding this protein, and those shown in T04234 and T04236. Histamine
 CC H1 receptor is part of a family of histamine receptor proteins. Of these
 CC sequences the greatest homology was found with bovine histamine H1
 CC receptor (overall identity of 82%, with the transmembrane regions alone
 CC having an identity of 96%). These receptors are coupled to different
 CC second messenger pathways via guanine nucleotide regulatory proteins
 CC (G proteins). This receptor (human histamine H1) transduces the signal
 CC through calcium ion mobilisation via an increase in intracellular
 CC inositol 1,4,5-triphosphate levels. Plasmids containing human histamine
 CC H1 receptor cDNA (pref. pNIV3604B) are used to transform mammalian cells
 CC (pref. CHO cells). These transformed cells express the receptor on the
 CC cell surface. The transformed cells can then be used to identify ligands
 CC that bind to the encoded protein, esp. for drug screening.

XX Sequence 330 AA;

Query Match 10.7%; Score 70.5; DB 16; Length 330;
 Best Local Similarity 30.7%; Pred. No. 3.2;
 Matches 23; Conservative 12; Mismatches 25; Indels 15; Gaps 4;

QY 38 VKGCEH-----AFAFCQKAKLKPSNP-GNNKFTIIDLVAAQLRRRLPARRGGKKQKHI 91
 Db 67 vrqhcqhrelnrlslpsfseikrlpenkpgdakkpkespwevlkrkpkdagggs---vl 123

QY 92 KPCSDSYEKRTPK 106
 Db 124 ksp-----qtpke 132

RESULT 15

R79965
 ID R79965 standard; Protein; 487 AA.
 XX

AC R79965;
 XX

DT 11-JAN-1996 (first entry)
 XX

DE Human histamine H1 receptor.
 XX

KW Histamine H1 receptor; human; bovine; transformation; drug screening;

KW pNIV3604B; calcium ion mobilisation; inositol 1,4,5-triphosphate;
 KW CHO cell; guanine nucleotide regulatory protein; G protein;
 KW hybridisation.
 XX

OS Homo sapiens.

Key	Location/Qualifiers
FT Domain	30..50 "transmembrane domain i"
FT Domain	/note= "transmembrane domain i"
FT Domain	69..89 "transmembrane domain ii"
FT Domain	/note= "transmembrane domain ii"
FT Domain	101..121 "transmembrane domain iii"
FT Domain	/note= "transmembrane domain iii"
FT Domain	145..165 "transmembrane domain iv"
FT Domain	/note= "transmembrane domain iv"
FT Domain	189..209 "transmembrane domain v"
FT Domain	/note= "transmembrane domain v"
FT Domain	419..439 "transmembrane domain vi"
FT Domain	/note= "transmembrane domain vi"
FT Domain	451..471 "transmembrane domain vii"
FT Domain	/note= "transmembrane domain vii"

GB2283239-A.

03-MAY-1995.

29-OCT-1993; 93GB-0022353.

29-OCT-1993; 93GB-0022353.

(UNIO) UCB SA.

Bollen A, Gillard M, Moguilevsky N, Noyer M, Varsalona F;

WPI: 1995-157301/21.

N-PSDB; T04227.

PT New nucleic acid encoding human histamine H1 receptor - useful
 PT diagnostically and for screening receptor binding drugs

PS Claim 6; Page 31-35; 49pp; English.

CC This sequence represents the protein for human histamine H1 receptor
 CC (constructed using T04234-6 and R79966-8). Histamine H1 receptor is part
 CC of a family of histamine receptor proteins. Of these sequences the
 CC greatest homology was found with bovine histamine H1 receptor (overall
 CC identity of 82%, with the transmembrane regions alone having an identity
 CC of 98%). These receptors are coupled to different second messenger
 CC pathways via guanine nucleotide regulatory proteins (G proteins). This
 CC receptor (human histamine H1) transduces the signal through calcium ion
 CC mobilisation via an increase in intracellular inositol 1,4,5-triphosphate
 CC levels. Plasmids containing human histamine H1 receptor cDNA (pref.
 CC pNIV3604B) are used to transform mammalian cells (pref. CHO cells).
 CC These transformed cells express the receptor on the cell surface. The
 CC transformed cells can then be used to identify ligands that bind to the
 CC encoded protein, esp. for drug screening.

XX Sequence 487 AA;

Query Match 10.7%; Score 70.5; DB 16; Length 487;
 Best Local Similarity 30.7%; Pred. No. 5.2;
 Matches 23; Conservative 12; Mismatches 25; Indels 15; Gaps 4;

QY 38 VKGCHCH- ----AAFACFQKAKLPSPN-GNNKTFIIDLVLAQLRRRLPARGGKKQKHIA 91

Db 217 vrqdcqhrelinslpsfseiklrpenpkgdakpkespwevkrpkpdkaggs---vl 273

QY 92 KPCSCDSYEKRTPE 106

Db 274 kps-----qtpke 282

Search completed: May 23, 2001, 11:42:44
 Job time: 1161 sec

```

Db      891 RILEKOGHLYVEVNGLEAYEAIKRNKYDVVLMDVQMPVNG-----GFATEKIRQWEKK 943
Qy      58 SNPNNKTFIDLVAQLRRRLPARRGGKKOKHIAKPCSDSYEKRTPKFEFLERLKWLLQK 117
Db      946 SNPIDSITFTPIIALTAHML-----GDREKSLAK--GMDDYVSRPLKPKL--LMOTINK 997
Qy      118 MH 120
Db      998 CIH:1000

RESULT 2
US-08-843-530B-33
: Sequence 33, Application US/08843530B
: Patent No. 5939306
: GENERAL INFORMATION:
: APPLICANT: Selitrennikoff, Claude
: APPLICANT: Agnan, Jacqueline
: APPLICANT: Alex, Lisa A.
: APPLICANT: Simon, Melvin I.
: TITLE OF INVENTION: Osmosensing Histidine Kinases
: NUMBER OF SEQUENCES: 36
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Medlen & Carroll, LLP
: STREET: 220 Montgomery Street, Suite 2200
: CITY: San Francisco
: STATE: California
: COUNTRY: United States of America
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/843,530B
: FILING DATE: 16-APR-1997
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Macknight, Kamrin T.
: REGISTRATION NUMBER: 38,230
: REFERENCE/DOCKET NUMBER: UTC-02717
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 705-8410
: TELEFAX: (415) 397-8338
: INFORMATION FOR SEQ ID NO: 33:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1117 amino acids
: TYPE: amino acid
: STRANDEDNESS: not relevant
: TOPOLOGY: not relevant
: MOLECULE TYPE: protein
US-08-843-530B-33

```

```
Query Match      12.1%; Score 80; DB 2; Length 1117;  
Best local Similarity 28.5%; Pred. No 0.31;  
Matches 35; Conservative 19; Mismatches 49; Indels 20; Gaps 7;  
  
QY   5    LRILRLRHLDIVEQ--LKIVE-----NDLPBELLSAPODVKGHCHEAAFCFAKAL--KP 57  
         ||| : |:::||| :|| :| :| :| :| :| :| :| :| :| :| :| :| :|  
Db    927 RILEKOGHLVEVNGLEAVEAIKRKYDVLMDVPQVMG-----GFATEKIRIQWEKK 981  
  
QY   58    SNPGNNTFTIIDLVAOLRRLLPARRGGKKOKHIAKCPSCDSYEKRTPEFLERLKWLLQK 117  
         ||| : |:::||| :||| :| :| :| :| :| :| :| :| :| :| :| :|  
Db    982 SNPIDSLTFRTPILATAHAML----GDREKS LAK---GMDDYVSVKPLPKKL--LMQTINK 1033  
  
QY   118 MIH 120  
        ||  
Db    1034 CIH 1036
```

RESULT 3

```

US-08-031-399-3
: Sequence 3, Application US/08031399
: Patent No. 5552303
: GENERAL INFORMATION:
: APPLICANT: Grabstein, Kenneth
: APPLICANT: Anderson, Dirk
: APPLICANT: Eisenman, June
: APPLICANT: Fung, Victor
: APPLICANT: Rauch, Charles
: TITLE OF INVENTION: Epithelium-derived T-cell Factor
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Immunex Corporation
: STREET: 51 University Street
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98101
: * COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/031.399
: FILING DATE: 19930308
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Launer, Charlene
: REGISTRATION NUMBER: 33,035
: REFERENCE/DOCKET NUMBER: 2811
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 206-587-0430
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 114 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHETICAL: NO
: US-08-031-399-3

Query Match 10.4%; Score 68.5; DB 1; Length 114;
Best Local Similarity 22.8%; Pred. No. 0.42;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps

Qy 10 LRHLIDIVEQLKIYENDLDPELLSAPQVKGHCEHAFAFCQAKLKPSNPGNNKTFIID 69
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 LKKIEDLIQSMHI-----DATLYTSDVHPCKVTAMKCF-LLEQLVISHESGDTDIHD 61
: : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 70 LVAQ---LRRRLPARGGKQKHIAKCSQSDSYKARTPKFELERLKWLLQKMIH 120
: | : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 TVENLIILANNILSNGNITE---SGCKECELEEKNIKEFLQSFVHIVQMFIN 112
: | : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
US-08-393-305-6
: Sequence 6, Application US/08393305
: Patent No. 5574138
: GENERAL INFORMATION:
: APPLICANT: Grabstein, Kenneth
: APPLICANT: Anderson, Dirk
: APPLICANT: Eisenman, June
: APPLICANT: Fung, Victor
: APPLICANT: Rauch, Charles
: TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Seed and Berry
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: Washington

```

QY	10	LRHLIDIVEOLKIYENDLDPELLSAPODVKGCHCEHAAFACQKAKLPSNPGNKFTIID	69
Db	9	LKKIEDLQSHHI-----DATLYESDVHPSCKVTAMKCF--LLEQLVISHESGDTIHD	61
QY	70	LVAQ---LRRRLPARRGGKKHIAKPCSCSYSEKRTPEFLERLKLWQMIH	120
Db	62	TVENLIILANILSSNGNITE---SGKCECELEEKNIKEFLQSFVHIVQFIN	112

ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/504,042
FILING DATE: 19-JUL-1995
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/031,399
FILING DATE: 08-MAR-1993

ATTORNEY/AGENT INFORMATION:
NAME: Launer, Charlene
REGISTRATION NUMBER: 33,035
REFERENCE/DOCKET NUMBER: 2811

TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-587-0430

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 114 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
HYPOTHETICAL: NO

55-08-504-042-3

Query Match 10 4% Score 68 5. DB 1. Length 114.

```

; Sequence 6, Application US/08794524
; Patent No. 5985262
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08794,524
; FILING DATE: 03-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,305
; FILING DATE: 22-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-794-524-6

Query Match      10.4%; Score 68.5; DB 2; Length 114;
Best Local Similarity 22.8%; Pred. No. 0.42;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps 69

QY   10 LRLHLDIVQEQLKIYENDLDPILLSAPQDVKGHCHEAAFCQKAKLKPSNPGNNKTFFID 69
    | : | ::::: | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    9 LKKEDTLQSMMHI-----DATLTESDVHSCKVTAMKCF--LLELQVISHESGDTDHD 61

QY   70 LVQA---LRRRLPARRGGKOKHIAKPCSDPSYEKRTPKEFLERLKLWLQKMIH 120
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    62 TVENLIILANLLSSNGNITE--SGCKECELEBKNIKEFLQSFVHVQMFIN 112


RESULT          9
US-09-189-193-6
; Sequence 6, Application US/09189193
; Patent No. 6184359
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle

```

```
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/189,193
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,305
; FILING DATE: 22-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-189-193-6

Query Match          10.4%; Score 68.5; DB 4; Length 114;
Best Local Similarity 22.8%; Pred. No. 0.42;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps 4;

QY 10 LRLHLDIVQKLYENDLPDLLSAPQDVKGHCHEAFAACFQKAKLKPSNPGNKKTFIID 69
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 9 LKKIEDLIQSMHI-----DATLYTESDVHPSCVKVTAMKCF-LLELQVISHESGDTDIHD 61

QY 70 LVAQ---LRRRLPARRGGKKQKHIAKPCSDSYEKRTPKFEFLERLKLWLLQKMIH 120
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 62 TVENLIILANNILSSNGNITE---SGCKECELEEKNIKEFLQSFVHIVQMFN 112

RESULT 10
PCT-US94-03793-3
; Sequence 3, Application PC/TUS9403793
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Interleukin-15
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03793
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
```

```
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; PCT-US94-03793-3

Query Match          10.4%; Score 68.5; DB 5; Length 114;
Best Local Similarity 22.8%; Pred. No. 0.42;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps 4;

QY 10 LRLHLDIVQKLYENDLPDLLSAPQDVKGHCHEAFAACFQKAKLKPSNPGNKKTFIID 69
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 9 LKKIEDLIQSMHI-----DATLYTESDVHPSCVKVTAMKCF-LLELQVISHESGDTDIHD 61

QY 70 LVAQ---LRRRLPARRGGKKQKHIAKPCSDSYEKRTPKFEFLERLKLWLLQKMIH 120
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 62 TVENLIILANNILSSNGNITE---SGCKECELEEKNIKEFLQSFVHIVQMFN 112

RESULT 11
US-08-300-903A-3
; Sequence 3, Application US/08300903A
; Patent No. 5591630
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M
; APPLICANT: Giri, Judith G
; TITLE OF INVENTION: Interleukin-15 Receptors
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESS: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple Operating System 7.1
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/300,903A
; FILING DATE: 06-SEPTEMBER-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/236,919
; FILING DATE: 06-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2822-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; TELEFAX: 206-233-0644
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-300-903A-3

Query Match          10.4%; Score 68.5; DB 1; Length 122;
```

```

Best Local Similarity 22.8%; Pred. No. 0.46;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps 4;

QY 10 LRHLIDIVEQLKIYENDLDPELLSAPODVKGCHCEHAAFACFOKAKLKPSNPGNNKFTIID 69
   | : : : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 17 LKKTIEDLIQSHI-----DATLYTESDVHPSCVKTKMCKF-LLEQLQVISHESGDTDHD 69
   | : : : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 70 LVAO---LRRRLPARRGGKKKHIAKPCSDSYEKRTPEFLERLKWLLQKMIH 120
   | : : : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 70 TVENLIILANNILSSNGNITE---SGCKECELEEKNIKEFLQSFVHVQMFIN 120
   | : : : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
US-08-031-399-2
; Sequence 2, Application US/08031399
; Patent No. 5552303
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Epithelium-derived T-cell Factor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/031,399
; FILING DATE: 19930308
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-031-399-2

Query Match 10.4%; Score 68.5; DB 1; Length 162;
Best Local Similarity 22.8%; Pred. No. 0.67;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps 4;

QY 10 LRHLIDIVEQLKIYENDLDPELLSAPODVKGCHCEHAAFACFOKAKLKPSNPGNNKFTIID 69
   | : : : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 57 LKKTIEDLIQSHI-----DATLYTESDVHPSCVKTKMCKF-LLEQLQVISHESGDTDHD 109
   | : : : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 70 LVAO---LRRRLPARRGGKKKHIAKPCSDSYEKRTPEFLERLKWLLQKMIH 120
   | : : : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 110 TVENLIILANNILSSNGNITE---SGCKECELEEKNIKEFLQSFVHVQMFIN 160
   | : : : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
US-08-393-305-5
; Sequence 5, Application US/08393305
; Patent No. 5574138
; GENERAL INFORMATION:

```

US-08-284-393B-9

Db 110 TVENLIILANILSSNGNITE--SGCKECELEEKNIKEFLOSFVHVIVOMFIN 160

; ZIP: 98104
COMPUTER BBS
:

LENGTH: 11.5 INCHES

Copyright (c) 1993 - 2000 CompuGen Ltd.
GENCORE version 4.3

Result No.	Score	Query %		Length	DB	ID	Description
		Match	Length				
1	79.5	12.1	414	2	S55654	DNA polymerase pro	
2	77	11.7	404	2	T27106	hypothetical prote	
3	75	11.4	718	1	A69084	DNA topoisomerase	
4	74	11.2	1081	2	T18223	histidine kinase -	
5	74	11.2	1081	2	T18258	histidine kinase -	
6	74	11.2	1081	2	T14003	histidine kinase h	
7	74	11.2	1081	2	T14173	histidine kinase h	
8	73	11.1	156	2	T10140	moda.3 protein - p	
9	72.5	11.0	510	2	SJ0724	cytochrome P450 AL	
10	72.5	11.0	698	2	S25409	transcription fact	
11	71	10.8	409	2	E83992	ATP/GTP-binding pr	
12	70.5	10.7	487	2	JC2495	histamine H1 recep	
13	70.5	10.7	874	2	A64664	valine--tRNA ligas	
14	70	10.6	440	2	T47906	FUSCA PROTEIN FUS6	
15	69.5	10.5	1324	2	S51622	cut3 protein - fis	
16	69	10.5	450	2	S00950	hypothetical prote	
17	68.5	10.4	162	1	A53484	interleukin-15 pre	
18	68.5	10.4	293	2	T31146	probable DNA inver	
19	68.5	10.4	491	2	A41632	histamine H1 recep	
20	68	10.3	1127	1	GNWVD2	genome polyprotein	
21	68	10.3	2749	1	ACMSIT	inositol 1,4,5-tri	
22	67.5	10.2	1142	2	A45031	cysteine-rich fibr	
23	66.5	10.1	324	2	S54361	gag-related protei	
24	66.5	10.1	443	2	D82975	two-component sens	
25	66.5	10.1	472	2	T15700	hypothetical prote	
26	66.5	10.1	853	2	H70470	hypothetical prote	
27	66.5	10.1	1256	2	C71436	probable resistanc	
28	66	10.0	128	2	T32947	hypothetical prote	
29	66	10.0	1045	2	S60571	integrin alpha v c	


```

Query Match          11.7%; Score 77; DB 2; Length 404;
Best Local Similarity 23.9%; Pred. NO. 4.5;
Matches 26; Conservative 19; Mismatches 34; Indels 30; Gaps 5;

Qy 26 DLDPELLSAPOQDKGCEHAFAFCQAKLK---PSPNGNKKTFIIDLVAQLRRRLPARR 82
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 206 DLNGLDKRLDKVSGVASAKTIFDMPKVKKEDPDPEPSQ-----PSRK 251
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

Qy 83 GGGKKOKHIAKPCSDSYEKRTP----KEFLERLK-----WLLQKMI 119
   ||||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 252 SGKKRSR-SRSPAANKFKSKSALDEIKEMEERKKERNRKDYWMREGIV 299
   ||||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

RESULT 3
A69084
C:Species: Methanobacterium thermoautotrophicum (strain Delta H)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A69084
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
  Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.;
  ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; MUID:98037514
A:Accession: A69084
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-718 <MTH>
A:Cross-references: GB:AE000921; GB:AE000666; NID:g2622744; PIDN:AA886097.1; PID:g262275
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH1624
A:Start codon: GTG
C:Superfamily: DNA topoisomerase I topA

```

	Query Match	11.4%	Score 75;	DB 1;	Length 718;	
	Best Local Similarity	26.1%;	Pred. No. 12;			
	Matches	35; Conservative	16; Mismatches	35; Indels	48; Gaps	8;
Qy	4	DRLLRLLRHLI-DIVVQLKIYENDLPDLSAPOD--VKGHCEHAFAACFOKAKLKPSNP	60			
	: : :	:	:	:	:	
Db	561	DEAITEVRSILSDIEENLr---DIGKELYRAYODSRWVGEP-----ACGGKLVIKYIS-P	611			
	: : :	:	:	:	:	
Qy	61	GNNKTFFI-----IDLVLAQLRRR----LPARRGGRKK-	86			
	: : :	:	:	:	:	
Db	612	RNRSTFVCGSSYPDCRTVYSLPRGASVLKSICEKGCLPMISYGPRPQRACLDPKCGKKKS	671			
Qy	87	--QKHIAKCPCSDS	98			
	: : : :	:	:	:	:	
Db	672	EVEEYVGGKPECGS	685			

```

RESULT      4
Ti18223    histidine kinase - yeast (Candida albicans)
C/Species: Candida albicans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: Ti18223
R;Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, November 1998
A/Reference number: Z18831
A/Accession: Ti18223
A/Status: preliminary; translated from GB/EMBL/DBDJ
A/Molecule type: DNA
A/Residues: 1-1081 <BAR>
A/Cross-references: EMBL:AL033396; NID:el1340089; PID:el1340096; PID:CAA21950.1
C/Genetics:
A/Note: Ca35A5.07

```

```

Query Match          11.2%; Score 74; DB 2; Length 1081;
Best Local Similarity 27.6%; Pred. No. 23;
Matches 34; Conservative 19; Mismatches 50; Indels 20; Gaps 7;

Qy 5 RLLRLRLHLIDIVEQ--LKIYE----NDLDPELLSAPQDVKGHCHEAFACFQAKL--KP 57
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 891 RILEKQGHSEVVEVNGLEAVEAIAKRNKYDVVLMDVQMPVMG-----GFETEKIQWEKK 945
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy 58 SNPGNKNKTFIDILVAQLRRRLPARRGGKKKHIAKPCSCDSYEKRTPKFEFLERLKWLLQK 117
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 946 SNPIDSLTFTPIIALTAHAML----GDREKSLAK--GMDDVSVSKPLPKL--LQMTINK 997
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy 118 MTH 120
   ||
Db 998 CIH 1000

RESULT 5
T18258
histidine kinase - yeast (Candida albicans)
C:Species: Candida albicans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18258
R:Alex, L.A.; Korch, C.; Selitrennikoff, C.P.; Simon, M.I.
Proc.Natl.Acad.Sci. U.S.A. 95, 7089-7073, 1998
A:Title: COS1, a two-component histidine kinase that is involved in hyphal development
A:Reference number: Z18845; MUID:98284059
A:Accession: T18258
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1081 <ALE>
A:Cross-references: EMBL:U69886; NID:g3243088; PID:g3243089; PIDN:AAC23929.1
C:Genetics:
A:Gene: chik1

```

```

Query Match      11.2%; Score 74; DB 2; Length 1081;
Best Local Similarity 27.6%; Pred. No. 23;
Matches 34; Conservative 19; Mismatches 50; Indels 20; Gaps 7;

QY 5 RLLRLRLHLDIVQEQ-LKIYE----NDLDPELLSAPQDVKGHCHEHAAACFQAKL--KP 57
Db 891 RLEKOGHSEVVEVNGLEAYEAIKRNKYDVVYMDVQMPVMG-----GFEATEKIRQWEKK 945
QY 58 SNPNGNKTFTIDLVAQLRRRLPARRRGKKQKHIAKPCSDSYVEKRTPEFLERLKLWLLQK 117
Db 946 SNPIDSLTFTPIIALTAHAML-----GDREKSLAK--GMDDYVSKPLAPKL--LMQTINK 997
QY 118 MIH 120
Db 998 CIH 1000

RESULT 6
Ti4003
histidine kinase homolog - yeast (Candida albicans)
C:Species: Candida albicans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: Ti4003
R:Nagahashi, S.; Mio, T.; Ono, N.; Yamada-Okabe, T.; Arisawa, M.; Bussey, H.; Yamada-
Microbiology 144, 425-432, 1998
A:Title: Isolation of CasLm1 and CanIkl, the genes for osmosensing histidine kinase h
A:Reference number: Z12498; MUID:98154430
A:Accession: Ti4003
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1081 <NAG>
A:Cross-references: EMBL:AB006363; NID:g29111163; PIDN:BAA24952.1; PID:g29111164
C:Genetics:
A:Gene: NIK1
Query Match      11.2%; Score 74; DB 2; Length 1081;

```

Best Local Similarity 27.6%; Pred. No. 23; Mismatches 19; Indels 20; Gaps 7; Matches 34; Conservative 19; Mismatches 50; Indels 20; Gaps 7;

QY 5 LLRLRLRLHLDIVEQ-LKIYE-----NDLDPILLSAPQDVKGCHHAFAACFQAKL--KP 57
 I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
 Db 891 RILEKQGHSEVVENGLEAYEAIKRNKYDVVLMQVMPVMG-----GFATEKIRQWEKK 945

QY 58 SPPGNNKTFIIDLVAQLRRRLPARRGGKKQKHIAKPCSDSYEKRTPKFEFLERLKLQK 117
 I I I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
 Db 946 SPPGNNKTFIIDLVAQLRRRLPARRGGKKQKHIAKPCSDSYEKRTPKFEFLERLKLQK 117
 I I I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
 QY 118 MIH 120
 Db 998 CIH 1000

RESULT 7
 T14173
 histidine kinase homolog - yeast (Candida albicans)
 C:Species: Candida albicans
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
 C:Accession: T14173
 R:Srikantha, T.; Tsai, L.; Enger, L.; Highley, K.; Soll, D.R.
 Microbiology 144, 2715-2729, 1998
 A:Title: The two-component hybrid kinase regulator Can1K1 of Candida albicans.
 A:Reference number: Z17902; MUID:99018821
 A:Accession: T14173
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1081 <SRI>
 A:Cross-references: EMBL:AF029092; NID:g3851456; PID:g3851457; PIDN:AAC72284.1
 C:Genetics:
 A:Gene: NIK1

Query Match 11.2%; Score 74; DB 2; Length 1081;
 Best Local Similarity 27.6%; Pred. No. 23; Mismatches 19; Indels 20; Gaps 7; Matches 34; Conservative 19; Mismatches 50; Indels 20; Gaps 7;

QY 5 LLRLRLRLHLDIVEQ-LKIYE-----NDLDPILLSAPQDVKGCHHAFAACFQAKL--KP 57
 I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
 Db 891 RILEKQGHSEVVENGLEAYEAIKRNKYDVVLMQVMPVMG-----GFATEKIRQWEKK 945

QY 58 SPPGNNKTFIIDLVAQLRRRLPARRGGKKQKHIAKPCSDSYEKRTPKFEFLERLKLQK 117
 I I I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
 Db 946 SPPGNNKTFIIDLVAQLRRRLPARRGGKKQKHIAKPCSDSYEKRTPKFEFLERLKLQK 117
 I I I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
 QY 118 MIH 120
 Db 998 CIH 1000

RESULT 8
 T10140
 modA.3 protein - phage T4
 N:Alternate names: gp modA.3; protein dda.7
 C:Species: phage T4
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 23-Jul-1999
 C:Accession: T10140; JZ0009
 R:Frazier, M.W.; Mosig, G.
 Gene 88, 7-14, 1990
 A:Title: The bacteriophage T4 gene mth whose product inhibits late T4 gene expression in
 A:Reference number: JH0136; MUID:90255970
 A:Accession: T10140
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-156 <FRA>
 A:Cross-references: EMBL:M30001; NID:g1532074; PID:g1532082
 A:Experimental source: strain GT7
 R:Mosig, G.; Colowick, N.C.
 personal communication, 1993
 A:Reference number: JZ0005
 A:Accession: JZ0009

A:Molecule type: DNA
 A:Residues: 1-156 <MOS>
 A:Note: submitted to T4 project
 C:Genetics:
 A:Gene: modA.3; dda.7
 A:Map position: 13.369-13.864

Query Match 11.1%; Score 73; DB 2; Length 156;
 Best Local Similarity 29.6%; Pred. No. 4.3; Mismatches 21; Conservative 15; Mismatches 25; Indels 10; Gaps 3;

QY 3 PDRLRLRLHLDIVEQ-LKIYE-----NDLDPILLSAPQDVKGCHHAFAACFQAKL--PS 58
 I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
 Db 50 PENIILIKRHEDIL-----QNTDPSFSSSEALTIRGYKRAHEYGLDFEDDKVKLAS 103

QY 59 NPGNNKTFIID 69
 I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
 Db 104 QPSKSKRTFIIIE 114

RESULT 9
 JS0724
 cytochrome P450 ALK6-A, alkane-inducible - yeast (Candida maltosa)
 N:Contains: oxidoreductase (EC 1.-.-.-)
 C:Species: Candida maltosa
 C>Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 20-Jun-2000
 C:Accession: JS0724
 R:Ohkuma, M.
 submitted to JIPID, July 1992
 A:Reference number: JS0721
 A:Accession: JS0724
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-510 <OHK>
 A:Cross-references: DDBJ:D12718; NID:g2183354; PIDN:BAA02212.1; PID:g2183355
 C:Genetics:
 A:Gene: ALK6-A
 C:Superfamily: Candida cytochrome P450 52A3; cytochrome P450 homology
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase
 F:310-480/Domain: cytochrome P450 homology <P45>
 F:458/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 11.0%; Score 72.5; DB 2; Length 510;
 Best Local Similarity 23.1%; Pred. No. 16; Mismatches 28; Conservative 20; Mismatches 52; Indels 21; Gaps 4;

QY 6 LLRLRLRLHLDIVEQ-LKIYE-----NDLDPILLSAPQDVKGCHHAFAACFQAKLPSNPGNNKT 65
 I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
 Db 325 LPFELSHHENVWTTLKEVVDQSPFDVESITFETIQNCIDYLRWCLFESLRVNSVPFNSRT 384

QY 66 FIDLVLAQLRRRLPARRGGKKQKHIAKPCSD-----SYEKRTPKFEFLERLK 112
 I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
 Db 385 ANKDTI-----LP-RGGGDCSHPILVKGGQVLFPLVYASNRQEKYFGKRPBEFIPE-R 436

QY 113 W 113
 Db 437 W 437

RESULT 10
 S25409
 transcription factor znf6 - human
 C:Species: Homo sapiens (man)
 C>Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 01-Dec-2000
 C:Accession: S25409
 R:Lloyd, S.L.; Sargent, C.A.; Chalmers, J.; Lim, E.; Habeebu, S.S.M.; Affara, N.A.
 Nucleic Acids Res. 19, 4835-4841, 1991
 A:Title: An X-linked zinc finger gene mapping to Xq21.1-q21.3 closely related to ZFX
 A:Reference number: S25409; MUID:92020112
 A:Accession: S25409
 A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-698 <LLO>

A:Cross-references: EMBL:X56465; NID:g38027; PIDN:CAA39837.2; PID:g5304925

A:Note: the authors translated the codon CAG for residue 4 as Glu, GAA for residue 53 as Arg, AAG for residue 414 as His, AGT for residue 575 as Val, GAT for residue 576 as Ala, C:Genetics:

A:Gene: znf6

C:Superfamily: zinc finger protein 2FP-36; LIM metal-binding repeat homology

Query Match 11.0%; Score 72.5; DB 2; Length 698;

Best Local Similarity 26.8%; Pred. No. 21;

Matches 26; Conservative 12; Mismatches 32; Indels 27; Gaps 5;

Qy 35 PDVVGKH-----CEHAAFACFQAKLK---PSNPGNNKTFIIDLVAQ----- 73

Db 454 PSELKKHMRTHGEPYQOCYQICFRCADQSNLKTHIKSHGNLPLKYKCEHCPQAFGDERE 513

Qy 74 LRRRLPARRGGKKKHIAKPCSDSYEKTPKFEFLER 110

Db 514 LQRLDLFQGHKTH---QCPHCD--HKSTNSSDLKR 544

RESULT 11

ATP/GTP-binding protein (ImpB/MucB/SamB family) BH2741 [imported] - Bacillus halodurans

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000

C:Accession: E83992

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20263314

A:Accession: E83992

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-409 <STO>

A:Cross-references: GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BAB06460.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH2741

Query Match 10.8%; Score 71; DB 2; Length 409;

Best Local Similarity 23.9%; Pred. No. 18;

Matches 28; Conservative 18; Mismatches 33; Indels 38; Gaps 5;

Qy 10 LRHLIDIVQLIYENDLD---PELLSAPQDVKGHCHEAAAFACFQAKLKPSNPGNNKT 65

Db 92 LERFTDLIEPFSIDFQFMDVTHSQRLFGAPRETAQKVQQAIV-----HET 136

Qy 66 FTIDLVAQLRRRLPARRGGKKKHIAKPCSDSYEKTRTP-----KEFLERLKWLL 115

Db 137 GV-----RARIIGMESKVLAKM-ACDNFAKMPSGVFLHYKERNELLWPL 181

RESULT 12

JC2495

histamine H1 receptor - human

C:Species: Homo sapiens (man)

C>Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 21-Jul-2000

C:Accession: JC2495; JC2035; S48144

R:Fukui, H.; Fujimoto, K.; Mizuguchi, H.; Sakamoto, K.; Horio, Y.; Takai, S.; Yamada, K.

Biochem. Biophys. Res. Commun. 201, 894-901, 1994

A:Title: Molecular cloning of the human histamine H1 receptor gene.

A:Reference number: JC2495; MUID:94271250

A:Accession: JC2495

A:Molecule type: DNA

A:Residues: 1-487 <FUK>

A:Cross-references: DBJ:D14436; NID:g506335; PIDN:BAA03319.1; PID:g538229

R:De Backer, M.D.; Gommeren, W.; Moereels, H.; Nobels, G.; Van Gompel, P.; Leysen, J.E.;

Biochem. Biophys. Res. Commun. 197, 1601-1608, 1993

A:Title: Genomic cloning, heterologous expression and pharmacological characterization of

A:Reference number: JC2035; MUID:94107375

A:Accession: JC2035

A:Molecule type: DNA

A:Residues: 1-487 <DB>

A:Cross-references: GB:X76786; NID:g442517; PIDN:CAA54182.1; PID:g442518

R:Mogullevsky, N.; Varsalona, F.; Noyer, M.; Gillard, M.; Guillaume, J.P.; Garcia, L.

Eur. J. Biochem. 224, 489-495, 1994

A:Title: Stable expression of human H(1)-histamine-receptor cDNA in Chinese hamster o

localisation of the gene.

A:Reference number: S48144; MUID:95010026

A:Accession: S48144

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-487 <MOG>

A:Cross-references: EMBL:234897; NID:g510295; PIDN:CAA84380.1; PID:g510296

C:Comment: This receptor mediates the increase in capillary permeability through immu

C:Genetics:

A:Gene: GDB:HRH1

A:Cross-references: GDB:303929; OMIM:600167

A:Map position: 3p21-3p14

C:Superfamily: vertebrate rhodopsin

C:Keywords: glycoprotein; phosphoprotein; receptor; transmembrane protein

F:30-49/Domain: transmembrane #status predicted <TM1>

F:64-83/Domain: transmembrane #status predicted <TM2>

F:102-121/Domain: transmembrane #status predicted <TM3>

F:136-165/Domain: transmembrane #status predicted <TM4>

F:190-209/Domain: transmembrane #status predicted <TM5>

F:419-438/Domain: transmembrane #status predicted <TM6>

F:451-470/Domain: transmembrane #status predicted <TM7>

F:518/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:107,194,198/Binding site: histamine (Asp, Thr, Asn) #status predicted

F:128,235,255,378,396,398/Binding site: phosphate (Ser) (covalent) (by protein kinase

F:140,142,390,478/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #sta

Query Match 10.7%; Score 70.5; DB 2; Length 487;

Best Local Similarity 30.7%; Pred. No. 23;

Matches 23; Conservative 12; Mismatches 25; Indels 15; Gaps 4;

Qy 38 VKGHCHEH----AAAFACFQAKLKPSNP-GNNKTFIIDLVAQLRRRLPARRGGKKKHIA 91

Db 217 VRQHCQREHLNRLSPSESEIKLRPENPKGDAKPKGESPWEVLARKPKDAGGGS---VL 273

Qy 92 KPCSDDSYEKTRPKE 106

Db 274 KSPS-----QTPKE 282

RESULT 13

A64664

valine--trna ligase (EC 6.1.1.9) - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 16-Jul-1999

C:Accession: A64664

R:Tombs, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKe

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey,

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467

A:Accession: A64664

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-874 <FOM>

A:Cross-references: GB:AE000621; GB:AE000511; NID:g2314301; PIDN:AAD08195.1; PID:g231

C:Superfamily: valine--trna ligase

C:Keywords: ligase

Query Match

Best Local Similarity 10.7%; Score 70.5; DB 2; Length 874;

Matches 30; Conservative 21; Mismatches 41; Indels 31; Gaps 5;

```

QY      8  IRLPHLIDIVEQLKIYENDL-----DPELLSAPQDVKGCHCEHAAPACQKAKLKPS  58
      :  | : | : | | : | : | : | : | : | : | : | : | : | : | : | : |
Db     756  VGLUREAIENTERLQTYAOKLARLEKYSVISSKPLKSVSDVGFCQ--TYANLENLDLSP-  812
      :  | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY     159  NPGNNKTFIIDVAQLRRRLRPARRGGKKQKHIAKCPSCDSYEKRTPKPEFLR---LKWL  114
      ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     813  -----LVARLKKQLEKEKLEKLLNLHN-----ENFVKNAPKSVLEKAKESLKL  857
      :  | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY     115  LOK 117
      :  | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     858  LEK 860

RESULT  14
T47906
FUSCA PROTEIN FUS6 - Arabidopsis thaliana
N:Alternate names: protein T20K12.40
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
R:De Haan, M.; Maarse, A.C.; Grivelli, L.A.; Mewes, H.W.; Lemcke, K.; Mayer,
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z24480
A:Accession: T47906
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-440 <DEH>
A:Cross-references: EMBL:AL1137898
A:Experimental source: Cultivar Columbia; BAC clone T20K12
C:Genetics:
A:Map position: 3
A:Introns: 130/3; 237/3; 278/1; 304/2; 382/3
A:Note: T20K12.40

```

Query Match	10.6%	Score 70;	DB 2;	Length 440;
Best Local Similarity	26.8%	Pred. No. 24;		
Matches	34;	Conservative	15;	Mismatches 42;
			Indels	36;
			Gaps	5;
Qy	10	LRHLIDIVEQLKIYE-----NLDLP	EL-----LSAPQDVKGHCEHAFACFQAKLKP	57
Db	223	LAHL-----ELKKYKLAARF	LDVNPGLNSYNVFIAPQDIATYGGICALASF	277
Qy	58	SNPGNNKTFIDLVAAQLRRRLP	ARRGGKKQKHIACPCSDSYEKRTPK--EFLERLKWLL	115
Db	278	FIDNINFNLELYPDVRELIN-----	DFYSRYASCLEYLASLKSNL	320
Qy	116	QKMTQHQ	122	
Db	321	LLDIHLH	327	

RESULT 15
S51622
cut3 protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 15-Jul-1995 #sequences_revision 01-Sep-1995 #text_change 17-Nov-2000
C:Accession: S51622; T39417
R:Saka, Y.; Sutan, T.; Yamashita, Y.; Saitoh, S.; Takeuchi, M.; Nakaseko, Y.; Yanagida, EMO J. 13, 4938-4952, 1994
A:Title: Fission yeast cut3 and cut14, members of a ubiquitous protein family, are required for normal cell cycle progression in fission yeast.
A:Reference number: S51622; MUID:95045386
A:Accession: S51622
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1324 <SAX>
A:Cross-references: EMBL:D30788; NID:g577659; PIDN:BAA06454.1; PID:g603501
R:McDougall, R.; Rajandream, M.A.; Barrrell, B.G.; Ramsperger, U.; Bothe, G.; Pohl, T.
submitted to the EMBL Data Library, July 1999
A:Reference number: 221853
A:Accession: T39417
A>Status: preliminary
A>Status: translated from GB/EMBL/DDBJ

A:Molecule type: DNA
A:Residues: 1-381, 'L', 383-392, 'F', 394-1315, 'I', 1317-1324 <MCD>
A:Cross-references: EMBL:AL096797; PIDN:CA846756.1; GSPDB:GN00067; SPDB:SPBC146.03c
A:Experimental source: strain 972h-; cosmid c146
C:Genetics:
A:Gene: SPDB:SPBC146.03c
A:Map position: 2
C:Superfamily: chromosome segregation protein SMC1

[illegible]

Search completed: May 23, 2001, 11:43:47
Job time: 1224 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:49:37 ; Search time 32.99 Seconds
(without alignments)
128.757 Million cell updates/sec

Title: US-09-522-217-56_COPY_23_146
Perfect score: 659
Sequence: 1 QGPDRLRLRLHLDIVEQL.....KEFLERLKLQMKIHOHLS 124

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76.5	11.6	162	1 IL15_FELCA	O97687 felis silve
2	73	11.1	156	1 Y01F_BPT4	P39425 bacterioph
3	72.5	11.0	510	1 CP5Q_CANMA	Q12587 candida mal
4	70.5	10.7	371	1 H2AY_HUMAN	O75367 homo sapien
5	70.5	10.7	487	1 HH1R_HUMAN	P35367 homo sapien
6	70.5	10.7	874	1 SYV_HELPY	P56000 helicobacte
7	69.5	10.5	1324	1 CUT3_SCHPO	P41004 schizosacch
8	69	10.5	450	1 V50K_BDVP	P09516 barley yell
9	68.5	10.4	162	1 IL15_BOVIN	Q28028 bos taurus
10	68.5	10.4	162	1 IL15_CERAE	P40221 cercopithe
11	68.5	10.4	162	1 IL15_MACMU	P48092 macaca mula
12	68.5	10.4	491	1 HH1R_BOVIN	P30546 bos taurus
13	68	10.3	1127	1 POLG_DEN2D	P30026 dengue viru
14	68	10.3	2749	1 IP3R_MOUSE	P11881 mus musculu
15	67.5	10.2	441	1 FUSC_ARATH	P45432 arabidopsis
16	67	10.2	132	1 IL4_CANFA	O77762 canis fami
17	66.5	10.1	443	1 PHOR_PSEAE	P23621 pseudomonas
18	66.5	10.1	689	1 KF1B_RAT	O88658 rattus norv
19	66.5	10.1	1270	1 DDX9_HUMAN	Q08211 homo sapien
20	66.5	10.1	1360	1 DDX3_MOUSE	O70133 mus musculu
21	65.5	9.9	417	1 PYRC_PYRHO	O57740 pyrococcus
22	65.5	9.9	872	1 SYV_HELPJ	Q92K61 helicobacte
23	65	9.9	228	1 MX1L_HUMAN	P50539 homo sapien
24	65	9.9	228	1 MX1L_MOUSE	P50540 mus musculu
25	65	9.9	335	1 SLAM_HUMAN	Q02874 rattus norv
26	65	9.9	370	1 H2AY_RAT	Q02874 rattus norv
27	64.5	9.8	162	1 IL15_HUMAN	P40933 homo sapien
28	64	9.7	351	1 YN1L_YEAST	P53860 saccharomyc
29	64	9.7	542	1 IMAL_SCHPO	O14063 schizosacch
30	64	9.7	1709	1 CHD1_HUMAN	Q14606 homo sapien
31	64	9.7	1711	1 CHD1_MOUSE	P40201 mus musculu
32	63.5	9.6	324	1 FE22_RAT	P97578 rattus norv
33	63.5	9.6	529	1 YEJF_ECOLI	P33916 escherichia

34	63	9.6	841	1 HS74_MOUSE	Q61316 mus musculu
35	62.5	9.5	112	1 Y059_METJA	Q60381 methanococc
36	62.5	9.5	313	1 TFS2_DROME	P20232 drosophila
37	62.5	9.5	318	1 NADA_SYNY3	P74578 synechocyst
38	62.5	9.5	329	1 HM40_CAEEL	Q19503 caenorhabdi
39	62.5	9.5	333	1 SUCA_HUMAN	P53597 homo sapien
40	62.5	9.5	333	1 SUCA_MOUSE	Q9WUM5 mus musculu
41	62.5	9.5	423	1 VG75_HSVII	Q00129 ictaluriid h
42	62.5	9.5	528	1 RAG2_CHICK	P25022 gallus gall
43	62.5	9.5	559	1 PRIM_HELPJ	Q9ZM49 helicobacte
44	62.5	9.5	679	1 TKT1_YEAST	P23254 saccharomyc
45	62.5	9.5	1043	1 TCF8_MESAU	Q60542 mesocricetu

ALIGNMENTS

RESULT 1

IL15_FELCA

ID IL15_FELCA STANDARD; PRT; 162 AA.

AC O97687;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE INTERLEUKIN-15 PRECURSOR (IL-15).

GN IL15.

OS Felis silvestris catus (Cat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

OX NCBI_TaxID=9685;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Lymph node;

RA Barger A.B., Dean G.A., Lavoy A.S.;

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

CC FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-

CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15

CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R

CC GAMMA BUT NOT IL-2R ALPHA (BY SIMILARITY).

CC SUBCELLULAR LOCATION: SECRETED.

CC SIMILARITY: BELONGS TO THE IL-15 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC EMBL; AF108148; AAD05268.1; -

DR Cytokine; Glycoprotein; Signal.

FT SIGNAL 1 29 POTENTIAL.

FT PROPEP 30 48 POTENTIAL.

FT CHAIN 49 162 INTERLEUKIN-15.

FT DISULFD 83 133 POTENTIAL.

FT DISULFD 90 136 POTENTIAL.

FT CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 127 127 N-LINKED (GLCNAC. .) (POTENTIAL).

SQ SEQUENCE 162 AA; 18412 MW; D8C7CE7F40110DD CRC64;

Query Match 11.6%; Score 76.5; DB 1; Length 162;
Best Local Similarity 24.6%; Pred. No. 0.45;
Matches 29; Conservative 19; Mismatches 49; Indels 21; Gaps 5;

Oy 15 DIVEOLKIVENDLD---PELLSAPQDVKGHCHEAFAACF---QKAKLKPSNPGNKKTF 66

| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

Db 52 DVISDLKIIDKIIQSLHIDATLTSTEDVHPNCKVTKAMKCFLLLELHVLSLESNKTHQV 111

| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

Oy 67 --IIDLVAQLRRRLPARRGGKQKHIK--CPSCDSYKRTPKFEFLERLKLWLLQKMIH 120

| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

Db 112 ENIIIL-----ANSLSSNRNITETGCKECELEEKNIKEFLQSPFHVHVFQFIN 160

```

RESULT 2
ID Y01F_BPT4          STANDARD;          PRT;          156 AA.
AC P39425;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 18.3 KDA PROTEIN IN MODB-MRH INTERGENIC REGION.
GN Y01F OR DDA.7.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae;
OC T4-like phages.
OX NCBI_TaxID=10665;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90255970; PubMed=1692800;
RA Frazier M.W., Mosig G.;
RT "The bacteriophage T4 gene mrh whose product inhibits late T4 gene
RT expression in an Escherichia coli rpoH (sigma 32) mutant.";
RL Gene 88:7-14(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M30001; AA07799.1;
KW Hypothetical protein.
SQ SEQUENCE 156 AA; 18331 MW; 7AB2F4957156D089 CRC64;

Query Match 11.1%; Score 73; DB 1; Length 156;
Best Local Similarity 29.6%; Pred. No. 1;
Matches 21; Conservative 15; Mismatches 25; Indels 10; Gaps 3;

OY 3 PDLRLRLHLIDIVLEQLKIYENDLPDLSAPQDVKGH---CEHAFAFCQKAKLK-PS 58
DB 50 PNIIILIKHTEDIL-----QNTDSPFSSEALTIGYKRAHEYGLEDFMDKVKLAS 103
OY 59 NPGNKKTFIID 69
DB 104 QPSKSKTFIIE 114

RESULT 3
ID CP50_CANNA          STANDARD;          PRT;          510 AA.
AC Q12587;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CYTOCHROME P450 52C2 (EC 1.14.14.-) (CYPLIIC2) (ALKANE-INDUCIBLE
DE P450-ALK6-A).
GN CYP52C2.
OS Candida maltosa (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5479;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95169275; PubMed=7865134;
RA Okuma M., Muraoka S., Tanimoto T., Fujii M., Ohta A., Takagi M.;
RT "CYP52 (cytochrome P450alk) multigene family in Candida maltosa:
RT identification and characterization of eight members.";
RL DNA Cell Biol. 14:163-173(1995).
CC -!- FUNCTION: TOGETHER WITH AN NADPH CYTOCHROME P450 THE ENZYME SYSTEM
CC CATALYZES THE TERMINAL HYDROXYLATION AS THE FIRST STEP IN THE
CC ASSIMILATION OF ALKANES AND FATTY ACIDS.

```

```

CC -!- CATALYTIC ACTIVITY: HYDROXYLATION OF N-ALKANES AT THE TERMINAL
CC POSITION.
CC -!- INDUCTION: BY VARIOUS ALKANES.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D12718; BAA02212.1;
DR InterPro; IPR001128;
DR InterPro; IPR002974;
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01239; EP450IICYP52.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme.
FT BINDING 458 458 HEME (BY SIMILARITY).
SQ SEQUENCE 510 AA; 58113 MW; 2AD3F6B48AC596B8 CRC64;

Query Match 11.0%; Score 72.5; DB 1; Length 510;
Best Local Similarity 23.1%; Pred. No. 4.4;
Matches 28; Conservative 20; Mismatches 52; Indels 21; Gaps 4;

OY 6 LLILRLHLIDIVEOLKIYENDLPDLSAPQDVKGHCEHAFAFCQKAKLPSNPGNKT 65
DB 325 LFFELSHENVTTLKEVVDQSFDPVESITETQNC DYLRWCLFESLRVNPSPFNSRT 384
OY 66 FIDLVQLRRRLPARGGKKQKIIAKPCSD-----SYEKRPKFEFLERLK 112
DB 385 ANKDTI-----LP-RGGEDCSHPILVKKGDVLFPLVYASNRQKYFGRAPEEPIE-R 436
OY 113 W 113
DB 437 W 437

RESULT 4
ID H2AY_HUMAN          STANDARD;          PRT;          371 AA.
AC O75367; O75377; Q9UP96;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CORE HISTONE MACRO-H2A.1 (HISTONE MACROH2A1) (MH2A1) (H2A.Y) (H2A/Y).
GN H2AFY OR MACROH2A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Liver;
RX MEDLINE=98390273; PubMed=9714746;
RA Lee Y., Hong M., Kim J.W., Hong Y.M., Choe Y.-K., Chang S.Y.,
RA Lee K.S., Choe I.S.;
RT "Isolation of cDNA clones encoding human histone macroH2A1 subtypes.";
RL Biochim. Biophys. Acta 1399:73-77(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=98318631; PubMed=9653160;
RA Mao M., Fu G., Wu J.-S., Zhang Q.-H., Zhou J., Kan L.-X., Huang Q.-H.,
RA He K.-L., Gu B.-W., Han Z.-G., Shen Y., Gu J., Yu Y.-P., Xu S.-H.,
RA Wang Y.-X., Chen S.-J., Chen Z.;
RT "Identification of genes expressed in human CD34(+) hematopoietic
RT stem/progenitor cells by expressed sequence tags and efficient full-
RT length cDNA cloning.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:8175-8180(1998).

```



```

FT BINDING      536      536      ATP (BY SIMILARITY).
SQ SEQUENCE    874 AA; 101360 MW; BB5668E2ACD5AD28 CRC64;

Query Match      10.7%; Score 70.5; DB 1; Length 874;
Best Local Similarity 24.4%; Pred. No. 13;
Matches 30; Conservative 21; Mismatches 41; Indels 31; Gaps 5;

QY      8  IRLRHLDIVLEOLKIYENDL-----DPELLSAPOQVKGHCHEAFAFCOKAKLKPS 58
DB      756  VGLREATENTERGTYAQKTLARLEKSVISSPKPKSVSDVGECQ--TYANLEMLDUSP- 812
QY      59  NPGNNKFTIIDVLQARRRRLPARRGGKKQKHIACPCSDSYSEKRTPKREFLER----LKWL 114
DB      813  -----LVARLKKQLEKLEKELKLNHN---ENPVKNAPKSVLEKAKESLKL 857
QY      115  LQK 117
DB      858  LEK 860

RESULT 7
CUT3_SCHPO STANDARD; PRT; 1324 AA.
AC P41004;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE CHROMOSOME SEGREGATION PROTEIN CUT3.
GN CUT3.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RX MEDLINE=95045386; PubMed=7957061;
RA Saka Y., Sutan T., Yamashita Y., Saitoh S., Takeuchi M.,
RA Nakaseko Y., Yanagida M.;
RT "Fission yeast cut3 and cut14, members of a ubiquitous protein
RT family, are required for chromosome condensation and segregation in
RT mitosis";
RL EMBO J. 13:4938-4952(1994).
CC -1- FUNCTION: REQUIRED FOR CHROMOSOME CONDENSATION AND SEGREGATION
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D30788; BAA06454.1; --
CC Mitosis; ATP-binding; Coiled coil; Nuclear protein.
CC NP_BIND 155 162
CC FT DOMAIN 310 337
CC FT DOMAIN 370 628
CC FT DOMAIN 825 1077
CC FT DOMAIN 1297 1324
CC FT DOMAIN 1324 AA; 150594 MW; 304DA5873291F837 CRC64;
CC -----
QY      1  OGPDRLLIRLRLDIVEQLKIYENDLPELLSAPOQVKGHCHEAFAFCOKAKLKPSN 59
DB      1  OGPDRLLIRLRLDIVEQLKIYENDLPELLSAPOQVKGHCHEAFAFCOKAKLKPSN 59

```

```

Db 283 EGDGGLLEYLIIIGTSKYKPIIENM--QELNSDDCAEKESRLKVLSEKAKLEDK 340
QY 60 PG-----NNKTFIDLVLAQLRRRLPARRGGKKOKKHIAK 92
Db 341 NSVLSFLKDNELPMQNOLYRTILYETRNKTKLVONLLNSVE-----GKLAHLEK 392
QY 93 CPSCD-----SVEKRTPKFEFLERLKLWLOK 117
Db 393 CEQTERDISEKNEVKSUREKAANKNDCTSEKTRQSTEQOTVK-IEQLKFLNKK 448
RESULT 8
V50K_BYDVP STANDARD; PRT; 450 AA.
AC P09516;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE 50 KDA PROTEIN (ORF 4)
OS Barley yellow dwarf virus (isolate PAV) (BYDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Luteoviridae;
OC Luteovirus.
OX NCBI_TaxID=12040;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88289355; PubMed=3399386;
RA Miller W.A., Waterhouse P.M., Gerlach W.L.;
RT "Sequence and organization of barley yellow dwarf virus genomic RNA.";
RL Nucleic Acids Res. 16:6097-6111(1988).
CC -1- SIMILARITY: TO ORF 6 OF BEET WESTERN YELLOW VIRUS, ORF6 OF
CC POTATO LEAFROLL VIRUS AND OF 54K PROTEIN OF BEET NECROSIS
CC YELLOW VEIN VIRUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07653; CAA30495.1; -
DR PIR; S00950; S00950.
DR InterPro; IPR002929; -
DR Pfam; PF01690; PLRV_ORF5; 1.
SQ SEQUENCE 450 AA; 49733 MW; 5878FA9361498205 CRC64;
-----
Query Match 10.5%; Score 69; DB 1; Length 450;
Best Local Similarity 27.3%; Pred. No. 8.7;
Matches 27; Conservative 11; Mismatches 31; Indels 30; Gaps 5;
QY 24 ENLDLP--ELLSAPQDVKGHC-----EHAFAFCQKAKLPSPNGNKKTFIDLVLAQLR 75
Db 322 EEDVLPSEQLSSPMDSGNIIIPKPEPEVLGTGQNIYPE-----DVPPMAR 371
QY 76 RRLPARRGGKKKHIAKPCSDSYEKRTPKF---FLERL 111
Db 372 QKL-----REANAPSTLLYERTPKKSGNLSRL 401
RESULT 9
IL15_BOVIN STANDARD; PRT; 162 AA.
AC Q28028;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

```

```

OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97426124; PubMed=9282828;
RA Canals A., Gasbarre L.C., Boyd P.C., Almeria S., Zarlega D.S.;
RT "Cloning and expression of bovine interleukin-15: analysis and
RT modulation of transcription by exogenous stimulation.";
RL J. Interferon Cytokine Res. 17:473-480(1997).
CC -1- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U42433; AA85130.1; -
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48 POTENTIAL.
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 127 127 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 162 AA; 18554 MW; 6A63CAA329EB8302 CRC64;
-----
Query Match 10.4%; Score 68.5; DB 1; Length 162;
Best Local Similarity 21.8%; Pred. No. 3;
Matches 24; Conservative 21; Mismatches 58; Indels 7; Gaps 3;
QY 16 IVEQLKIYEN-----DIDPELLSAPQDVKGHCHEAFAFCQKAKLPSPNGNKKTFIDL 70
Db 53 VINDLKTIEHLIQSIHMDATLYT-ESDAHPNCKVTAMQCFLELRLVILHESKNAT-IYEI 110
QY 71 VAQLRRRLPARRGGKKKHIAKPCSDSYEKRTPKFLERLKWLQMKH 120
Db 111 IENLTMANSNLSSNIENKTELGCCELEKEKSIKEFLKSFVHVQMFN 160
RESULT 10
IL15_CERAE STANDARD; PRT; 162 AA.
AC P40221;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 49-81.
RC TISSUE-Kidney.
RX MEDLINE=94233380; PubMed=8178155;
RA Grabstein K.H., Eisenman J., Shanbeck K., Rauch C.,
RA Srinivasan S., Fung V., Beers C., Richardson J., Schoenborn M.A.,
RA Ahliah M., Johnson L., Alderson M.R., Watson J.D., Anderson D.M.,

```

RA Giri J.G.;
RT Cloning of a T cell growth factor that interacts with the beta chain
of the interleukin-2 receptor.;
RL Science 264:965-968(1994).
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U03099; AAA18416.1; -
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48 INTERLEUKIN-15.
FT CHAIN 49 162
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 162 AA; 18222 MW; 1BP9A82644E1C9B7 CRC64;

Query Match 10.4%; Score 68.5; DB 1; Length 162;
Best Local Similarity 22.8%; Pred. No. 3;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps 4;

Qy 10 LRHLIDIVOLKIYENDLPPELLSAPQDVKGCHCEHAFAACFKAKLKPSNPGNKFTIID 69
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 57 LKKIEDLIQSMHI-----DATLYTESDVHPCKVTAMKCF-LLELQVISHESGDTDIHD 109

Qy 70 LVAQ---LRRLPARRGGKQKHIAKPCSDSYEKRTPKFELRLKWLQKMIH 120
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 110 TVENLIILANNILSSNGNITE---SGCKEELEEKNKEFLQSFVHVQMFN 160

RESULT 11
ID IL15_MACMU STANDARD; PRT; 162 AA.
AC P48092;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=96003435; PubMed=7561102;
RA Villingier F.J., Brar S.S., Wayne A.E., Chikkala N., Ansari A.A.;
RT "Comparative sequence analysis of cytokine genes from human and
RT nonhuman primates.";
RL J. Immunol. 155:3946-3954(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Tatsumi M.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA.

CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U19843; AAB60398.1; -
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48 INTERLEUKIN-15.
FT CHAIN 49 162
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 5 31 K -> T.
FT VARIANT 31 31 I -> T.
SQ SEQUENCE 162 AA; 18194 MW; D233CF7FF6188C01 CRC64;

Query Match 10.4%; Score 68.5; DB 1; Length 162;
Best Local Similarity 22.8%; Pred. No. 3;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps 4;

Qy 10 LRHLIDIVEQLKIYENDLPPELLSAPQDVKGCHCEHAFAACFKAKLKPSNPGNKFTIID 69
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 57 LKKIEDLIQSMHI-----DATLYTESDVHPCKVTAMKCF-LLELQVISHESGDTDIHD 109

Qy 70 LVAQ---LRRLPARRGGKQKHIAKPCSDSYEKRTPKFELRLKWLQKMIH 120
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 110 TVENLIILANNILSSNGNITE---SGCKEELEEKNKEFLQSFVHVQMFN 160

RESULT 12
ID HHIR_BOVIN STANDARD; PRT; 491 AA.
AC P30546;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE HISTAMINE H1 RECEPTOR.
GN HH1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adrenal medulla;
RX MEDLINE=92107981; PubMed=1722337;
RA Yamashita M., Fukui H., Sugama K., Horio Y., Ito S., Mizuguchi H.,
RA Wada H.;
RT "Expression cloning of a cDNA encoding the bovine histamine H1
RT receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:11515-11519(1991).
CC -!- FUNCTION: IN PERIPHERAL TISSUES, THE H1 SUBCLASS OF HISTAMINE
CC RECEPTORS MEDIATES THE CONTRACTION OF SMOOTH MUSCLES, INCREASE IN
CC CAPILLARY PERMEABILITY DUE TO CONTRACTION OF TERMINAL VENULES,
CC AND CATECHOLAMINE RELEASE FROM ADRENAL MEDULLA, AS WELL AS
CC MEDIATING NEUROTRANSMISSION IN THE CENTRAL NERVOUS SYSTEM.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: BRAIN, LUNG, SMALL INTESTINE, UTERUS, ADRENAL
CC MEDULLA AND SPLEEN.
CC -!- PTM: POTENTIAL SITES OF PHOSPHORYLATION IN THE THIRD CYTOPLASMIC
CC LOOP MAY PLAY AN IMPORTANT ROLE IN REGULATING SIGNAL TRANSDUCTION
CC THROUGH THE RECEPTOR MOLECULE.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; DI0197; BAA01045.1; -
CC PIR; A41632; A41632.
CC GCRDB; GCR_0278; -
CC InterPro; IPR000276; -
CC InterPro; IPR000921; -
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHHODOPSN.
CC PRINTS; PR00530; HISTAMINEHR.
CC PROSITE; PS00237; G_PROTEIN_RECP_FL_1; 1.
CC PROSITE; PS0262; G_PROTEIN_RECP_FL_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Lipoprotein; Palmitate; Phosphorylation.
CC KW DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 31 50 1 (POTENTIAL).
CC FT DOMAIN 51 64 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 65 84 ~2 (POTENTIAL).
CC FT DOMAIN 85 102 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 103 124 3 (POTENTIAL).
CC FT DOMAIN 125 146 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 147 166 4 (POTENTIAL).
CC FT DOMAIN 167 190 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 191 211 5 (POTENTIAL).
CC FT DOMAIN 212 222 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 423 442 6 (POTENTIAL).
CC FT DOMAIN 443 454 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 455 474 7 (POTENTIAL).
CC FT DOMAIN 475 491 CYTOPLASMIC (POTENTIAL).
CC FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 18 18 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT DISULFID 101 181 BY SIMILARITY.
CC FT LIPID 448 PALMITATE (POTENTIAL).
CC SEQUENCE 491 AA; 559597 MW; DAA349E52218CE28 CRC64;
SQ

```

Query Match          10.4%; Score 68.5; DB 1; Length 491;
Best Local Similarity 26.8%; Pred. No. 11;
Matches 22; Conservative 11; Mismatches 28; Indels 21; Gaps 4;

QY 38 VKGCEH-----AAFACFQKAKLKPSN-----PGNKKTFIDLVAQLRRRLPARGGK 85
   1: 11:1 :1:1 :1:11 :1:1 : : :1:1 :1:1
Db 218 VRHCOHRELINGSFFSFSOMKMKPNQLVGAKKPGKESPW-----EVLKRPKPDGTTGG 271
   1: 11:1 :1:1 :1:11 :1:1 : : :1:1 :1:1
QY 86 KQKHIAKPCDCSVYEKRTPTREF 107
   : 11:1 :1:1 :1:1
Db 272 P----VLKPPSQEPKEVTS PGVF 290
   : 11:1 :1:1 :1:1

RESULT 13
POLG_DEN2D
ID POLG_DEN2D STANDARD; PRT; 1127 AA.
AC F30026; Q66450;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX
DE PROTEIN (ENVELOPE GLYCOPROTEIN M); MAJOR ENVELOPE PROTEIN E;
DE NONSTRUCTURAL PROTEIN NS1 (FRAGMENT').
OS Dengue virus type 2 (strain D2-04).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OX NCBI_TaxID=31636;
RN [1]
RP SEQUENCE OF 1-775 FROM N.A.
RA Yang P.Y., Lam S.K.;

```

```

RT  "The nucleotide and encoded amino acid sequences of the structural
RT  protein gene of D2-04 virus strain isolated in China.";
RL  Chin. J. Microbiol. Immunol. 11:341-344(1991).
RN  [2]
RP  SEQUENCE OF 776-1127 FROM N.A.
RA  Yang P.Y., Kautner I.M., Koh C.L., Lam S.K.;
RT  "Nucleotide and encoded amino acid sequences of the nonstructural
RT  protein NS1 gene of a Dengue-2 virus isolated in China.";
RL  Chin. J. Microbiol. Immunol. 11:9-12(1991).
RN  [2]
RP  SUBUNIT. THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
RA  LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
RT  PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
RL  PROTEIN C AND MRNA.
RN  [2]
RP  This SWISS-PROT entry is copyright. It is produced through a collaboration
RA  between the Swiss Institute of Bioinformatics and the EMBL outstation -
RT  the European Bioinformatics Institute. There are no restrictions on its
RL  use by non-profit institutions as long as its content is in no way
RN  modified and this statement is not removed. Usage by and for commercial
RP  entities requires a license agreement (See http://www.isb-sib.ch/announce/
RA  or send an email to license@isb-sib.ch).
RN  [2]
RP  ENBL; U19778; AAA73471.1; -
RA  PIR; JC1007; GNWVD2.
RT  HSSP; P14336; ISVB.
RN  [2]
RP  InterPro: IPR000069; -
RA  InterPro: IPR000336; -
RT  InterPro: IPR001122; -
RN  InterPro: IPR001157; -
RA  InterPro: IPR002535; -
RT  Pfam; PF01004; Flavi_M; 1.
RN  Pfam; PF00948; Flavi_NS1; 1.
RT  Pfam; PF01003; Flavi_capsid; 1.
RN  Pfam; PF00869; Flavi_glycoprot; 1.
RT  Pfam; PF01570; Flavi_propept; 1.
RN  [2]
RP  Polyprotein; Glycoprotein; Core protein; Coat protein;
RA  Envelope protein; Transmembrane; Nonstructural protein.
RT  CAPSID PROTEIN C.
RN  [2]
RP  CHAIN 1 114
RT  PROPEP 115 205
RA  CHAIN 206 280
RT  CHAIN 281 775
RN  [2]
RP  CHAIN 776 1127
RT  TRANSMEM 101 117
RA  TRANSMEM 727 743
RT  TRANSMEM 757 773
RN  [2]
RP  DISULFID 283 310
RA  DISULFID 340 396
RT  DISULFID 354 385
RN  [2]
RP  DISULFID 372 401
RA  DISULFID 465 565
RT  DISULFID 582 613
RN  [2]
RP  CARBOHYD 183 183
RA  CARBOHYD 347 347
RT  CARBOHYD 433 433
RN  [2]
RP  CARBOHYD 905 905
RA  CARBOHYD 982 982
RT  NON_TER 1127 1127
RA  SEQUENCE 1127 AA; 125835 MW; C93D541432DA80EE CRC64;
RN  [2]
RP  Query Match 10.3%; Score 68; DB 1; Length 1127;
RA  Best Local Similarity 26.2%; Pred. No. 32;
RN  [2]
RP  Matches 21; Conservative 20; Mismatches 23; Indels 16; Gaps 3;
RA  [2]
RP  QY 5 RLILRLHLI--DIVEQLKTYENDLDELPSAPDQVKHGCEHA-----AFAC 49
RT  I : I I I : I : I I : : : : I : I I I :
RA  833 RSVTRLENLAWKQITPELKKHLSIEVVKLTIMTGDIKGIMQAGTRSLRPOPTKLKPSWET 892
RN  [2]
RP  QY 50 FQAKLAPSPNPGNKFTIID 69
RT  : : I I I : I : I : I I I I I I
RA  893 WRKAKMVPTEP-HNQTEIID 911
RN  [2]

```

```
Query Match      10.3%; Score 68; DB 1; Length 1127;
Best Local Similarity 26.2%; Pred. No. 32;
Matches 21; Conservative 20; Mismatches 23; Indels 16; Gaps 3;

QY      5    LLRIRLRHLI--DIVEOLKIYENDLPDLISAPODVKGHCHEA-----AFAC 49
          | : || ::| : || ::::| : || ::||| :
Db      833 RSVTRLENLAWKQITPELKHTLSIEVKLTIMTGDIKGIMQAGTSRLRPQTLEKFSWET 892

QY      50 FQAKLKPSPNGNKKFTIID 69
          ::||| :| :| :|||::|
Db      893 WRKAKMVPTEP-HNQTELID 911
```

RESULT 14
IP3R_MOUSE
ID IP3R_MOUSE STANDARD; PRT; 2749 AA.
AC P11881;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN TYPE 1 RECEPTOR
DE (INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN P(400)) (TYPE 1 INSP3
DE RECEPTOR).
GN INSP3R OR ITPR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Purkinje cells; PubMed=2554142;
RX MEDLINE=90044039; PubMed=2762133;
RA Furuichi T., Yoshikawa S., Miyawaki A., Wada K., Maeda N.,
RA Mikoshiba K.;
RT "Primary structure and functional expression of the inositol 1,4,5-
RT trisphosphate-binding protein P400.";
RL Nature 342:32-38(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Cerebellum;
RX MEDLINE=89345101; PubMed=2762133;
RA Furuichi T., Yoshikawa S., Mikoshiba K.;
RT "Nucleotide sequence of cDNA encoding P400 protein in the mouse
RT cerebellum.";
RL Nucleic Acids Res. 17:5385-5386(1989).
RN [3]
RP SEQUENCE OF 318-332 AND 1692-1731 FROM N.A., AND ALTERNATIVE SPLICING.
RC STRAIN=JCR;
RX MEDLINE=91296797; PubMed=1648733;
RA Nakagawa T., Okano H., Furuichi T., Aruga J., Mikoshiba K.;
RT "The subtypes of the mouse inositol 1,4,5-trisphosphate receptor are
RT expressed in a tissue-specific and developmentally specific manner.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:6244-6248(1991).
CC -1- FUNCTION: RECEPTOR FOR INOSITOL 1,4,5-TRISPHOSPHATE, A SECOND
CC MESSENGER THAT MEDIATES THE RELEASE OF INTRACELLULAR CALCIUM.
CC THE RECEPTOR CONTAINS A CALCIUM CHANNEL IN ITS C-TERMINAL
CC EXTREMITY. ITS LARGE N-TERMINAL CYTOPLASMIC REGION HAS THE LIGAND-
CC BINDING SITE IN THE N-TERMINUS AND MODULATORY SITES IN THE MIDDLE
CC PORTION IMMEDIATELY UPSTREAM OF THE CHANNEL REGION.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC
CC RETICULUM.
CC -1- ALTERNATIVE PRODUCTS: ADDITIONAL SUBTYPES OF INSP3R ARISE BY
CC ALTERNATIVE SPLICING OF THE SAME GENE.
CC -1- PTM: PHOSPHORYLATED BY CYCLIC-AMP KINASE. PHOSPHORYLATION PREVENTS
CC THE LIGAND-INDUCED OPENING OF THE CALCIUM CHANNELS.
CC -1- MISCELLANEOUS: CALCIUM APPEARS TO INHIBIT LIGAND BINDING TO THE
CC RECEPTOR, MOST PROBABLY BY INTERACTING WITH A DISTINCT
CC CALCIUM-BINDING PROTEIN WHICH THEN INHIBITS THE RECEPTOR.
CC -1- SIMILARITY: TO RYANODINE RECEPTOR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X15373; CAA33433.1; -
CC EMBL; M75986; AAA39316.1; -
CC EMBL; M75987; AAA39317.1; -
CC PIR; S04844; ACMSIT.
CC MGD; MGI:96623; Itp1.
CC InterPro; IPR000493; -
CC InterPro; IPR000699; -

DR PFAM: PF01365; RYDR_LTPR; 1.
DR PRINTS; PR00779; INSP3RECEPTR.
KW Receptor; Transmembrane; Glycoprotein; Phosphorylation;
KW Endoplasmic reticulum; Ionic channel; Ion transport; Calcium channel;
KW Alternative splicing.
FT DOMAIN 1 2273 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 2274 2294 M1 (POTENTIAL).
FT TRANSMEM 2308 2326 M2 (POTENTIAL).
FT TRANSMEM 2334 2356 M3 (POTENTIAL).
FT TRANSMEM 2365 2387 M4 (POTENTIAL).
FT TRANSMEM 2391 2407 M5 (POTENTIAL).
FT TRANSMEM 2440 2462 M6 (POTENTIAL).
FT TRANSMEM 2530 2549 M7 (POTENTIAL).
FT TRANSMEM 2570 2589 M8 (POTENTIAL).
FT DOMAIN 2590 2749 CYTOPLASMIC (POTENTIAL).
FT MOD_RES 1588 1588 PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT MOD_RES 1755 1755 PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT VARSPLIC 318 332 MISSING (IN ISOFORM INSP3R SI-).
FT VARSPLIC 1692 1731 MISSING (IN ISOFORM INSP3R SIABC-).
FT VARSPLIC 1715 1715 MISSING (IN ISOFORM INSP3R SIIB-).
FT VARSPLIC 1715 1731 MISSING (IN ISOFORM INSP3R SIIBC-).
SQ SEQUENCE 2749 AA; 313193 MW; 47E5F24BCD5F4153 CRC64;

Query Match 10.3%; Score 68; DB 1; Length 2749;
Best Local Similarity 28.1%; Pred. No. 89;
Matches 32; Conservative 14; Mismatches 48; Indels 20; Gaps 6;

QY 4 DRU--LIR--LRHLIDIVEOLKIYENDLDPILLSAPQ-DVKGHCEHAFAACFQAKLKPS 58
||| : : : : : ||| : : : : : ||| : : : : :
DB 1612 DRUPLVQAEISLVLDVLR-----PELLFPNTDARRKCESGGFICKLIKHTKQL 1662
QY 59 NPGNNKTFIDLVQLRRRLPARRG-GKKQKHI-----AKPCSDSYEKRTPK 106
: : : : : ||| : : : : : ||| : : : : :
DB 1663 LEENEKCKICIKVLQTLREMTKDRGYGEKQISIDSENAELPQAPAEENSTEQE 1716

RESULT 15
FUS6_ARATH
ID FUS6_ARATH STANDARD; PRT; 441 AA.
AC P45432;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FUSCA PROTEIN FUS6.
GN FUS6 OR COP1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. WASSILEWSKIJA; TISSUE=Siliques;
RX MEDLINE=94176998; PubMed=8130643;
RA Castle L.A., Meinke D.W.;
RT "A FUSCA gene of Arabidopsis encodes a novel protein essential for
RT plant development.";
RL Plant Cell 6:25-41(1994).
CC -1- FUNCTION: IS AN ESSENTIAL REGULATORY PROTEIN. HAS AN ESSENTIAL
CC ROLE THROUGHOUT PLANT DEVELOPMENT. IT COULD BE IMPORTANT IN
CC RELAYING PLANT HORMONE AND OTHER SIGNALS FROM MEMBRANE-ASSOCIATED
CC PKC TO CYTOPLASMIC FACTORS AND THIS COULD LEAD DIRECTLY TO
CC BIOCHEMICAL RESPONSES OR TO DOWNSTREAM GENE REGULATION.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC -1- TISSUE SPECIFICITY: LEAVES, FLOWERS, IMMATURE SILIQUES, AND
CC LIGHT-GROWN ROOTS.
CC -1- SIMILARITY: BELONGS TO THE FUS6 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: L26498; AAA32792.1; -.
DR DR InterPro: IPR000717; -.
DR DR Pfam: PF01399; PCI: 1.
DR DR
SQ SEQUENCE 441 AA; 50548 MW; 2F6F223FCC337756 CRC64;

Query Match	10.2%;	Score 67.5;	DB 1;	Length 441;
Best Local Similarity	27.3%;	Pred. No. 12;		
Matches 35;	Conservative 15;	Mismatches 41;	Indels 37;	Gaps 6;

QY	10	LRHLIDIVQLKTYE-----NDLDPEL-----LSAPQDVKGHCHEAAFCQKAKLKP	57
		: :	: :
Db	223	LAHL-----ELKKYLAARFLDVNPDELGSYNEVIAPQDIATYGGLCALASEFDSLELQ	277
		: :	: :
QY	58	SNPGN-NKTFIDLVQAQLRRRLPARRGGKKQKHIACPCSDSYEKKTPK--EFLERLKWL	114
		: :	: :
Db	278	KVIDNINFRNFLEVPDRELIN-----DFISSRYASCLEVLASELSKN	320
		: :	: :
QY	115	LQKMIHQH	122
		: :	: :
Db	321	LLLDIHLH	328
		: :	: :

Search completed: May 23, 2001, 11:49:38
Job time: 345 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:49:00 ; Search time 96.38 Seconds
(without alignments)
150.797 Million cell updates/sec

Title: US-09-522-217-56_COPY_23_146

Perfect score: 659

Sequence: 1 QGPDRLLRLRLHLDIVEQL.....KEFLERLKLWLLQKMHQHLS 124

Scoring table: BLOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP TREMBL15: *
2: sp_archaea: *
3: sp_bacteria: *
4: sp_fungi: *
5: sp_human: *
6: sp_invertebrate: *
7: sp_mammal: *
8: sp_mhc: *
9: sp_organelle: *
10: sp_phase: *
11: sp_plant: *
12: sp_rodent: *
13: sp_unclassified: *
14: sp_vertebrate: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80.5	12.2	513	Q9N8N9	Q9n8n9 trypanosoma
2	79.5	12.1	414	Q66661	Q66661 equine herp
3	78	11.8	1304	Q9SUK4	Q9suk4 arabidopsis
4	77	11.7	404	Q9XWF2	Q9xwf2 caenorhabdi
5	76	11.5	673	Q9SA08	Q9sa08 arabidopsis
6	75.5	11.5	372	Q93327	Q93327 gallus gall
7	75.5	11.5	494	Q9VWU0	Q9vwu0 drosophila
8	75	11.4	718	Q27661	Q27661 methanobact
9	75	11.4	1054	Q9W493	Q9w493 drosophila
10	74	11.2	354	Q65168	Q65168 african swi
11	74	11.2	566	Q86854	Q86854 streptomyce
12	74	11.2	797	Q9RKS9	Q9rks9 streptomyce
13	74	11.2	1081	Q42696	Q42696 candida alb
14	74	11.2	1081	Q74271	Q74271 candida alb
15	74	11.2	1081	Q93851	Q93851 candida alb
16	74	11.2	1081	Q9URL9	Q9url9 candida alb
17	73.5	11.2	203	Q9NN71	Q9nn71 leishmania
18	73.5	11.2	1361	O04264	O04264 arabidopsis
19	73	11.1	474	Q9N9T8	Q9n9t8 leishmania

20	72.5	11.0	372	11	Q9QZQ8	Q9qzq8 mus musculu
21	72.5	11.0	698	4	Q9Y4J6	Q9y4j6 homo sapien
22	72.5	11.0	771	4	Q9Y462	Q9y462 homo sapien
23	72	10.9	135	11	O88648	O88648 marmota mon
24	72	10.9	657	10	Q9ZUT5	Q9zut5 arabidopsis
25	71.5	10.8	487	6	Q9N2B2	Q9n2b2 pan troglod
26	71.5	10.8	487	6	Q9N2B1	Q9n2b1 gorilla gor
27	71	10.8	409	2	Q9K9A8	Q9k9a8 bacillus ha
28	70.5	10.7	223	13	O93325	O93325 gallus gall
29	70.5	10.7	369	13	O93326	O93326 gallus gall
30	70.5	10.7	372	4	O75367	O75367 homo sapien
31	70.5	10.7	487	6	Q9M2B0	Q9m2b0 pongo pygma
32	70	10.6	440	10	Q9M2E9	Q9m2e9 arabidopsis
33	69.5	10.5	162	6	Q9XSJ6	Q9xsj6 ovis aries
34	69	10.5	835	14	P88992	P88992 murid herpe
35	68.5	10.4	293	2	O85855	O85855 sphingomona
36	67.5	10.2	239	10	Q9LTG6	Q9ltg6 arabidopsis
37	67.5	10.2	1077	13	Q91019	Q91019 gallus gall
38	67.5	10.2	1142	13	Q02391	Q02391 gallus gall
39	67	10.2	371	4	Q9UP96	Q9up96 homo sapien
40	67	10.2	449	2	Q9S1J9	Q9s1j9 streptococc
41	67	10.2	451	4	O60319	O60319 homo sapien
42	67	10.2	528	4	O9NSF3	O9nsf3 homo sapien
43	67	10.2	617	4	Q9NRP6	Q9nrp6 homo sapien
44	67	10.2	818	4	Q9UP55	Q9ups5 homo sapien
45	66.5	10.1	305	5	O15561	O15561 nosema locu

ALIGNMENTS

RESULT 1

Q9N8N9 PRELIMINARY; PRT; 513 AA.
AC Q9N8N9:
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE PROBABLE AXONEME CENTRAL APPARATUS PROTEIN.
GN CHL1.235.
OS Trypanosoma brucei.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TREU927;
RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,
RA Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,
RA Gerrard C., Rajandream M.A., Barrrell B.G.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL359782; CAB95504.1; -
SQ SEQUENCE 513 AA; 56126 MW; 40BB4546A1B390C3 CRC64;

Query Match 12.2%; Score 80.5; DB 5; Length 513;
Best Local Similarity 23.5%; Pred. No. 2.8;
Matches 28; Conservative 24; Mismatches 52; Indels 15; Gaps 3;
QY 10 LRHLIDIVQLKIYENDLPDLSAPQDVKGHCHEAFAFCQKAKLKPSNPGNKKFTIID 69
Db 404 LKNIIRCVQLPALEPLLHPE---APKNV-----LKYYVCGQPAKVLPTDIAAKREFVAN 454
QY 70 LVAQLRRLLPARRGGCKKOKHIAKPCSC-----DSYEKRTPKPEFLERLKLWLLQKMHQH 122
Db 455 RGLATVQRIQPEPGSKLAETIOSINNCYPPEIVQVYSPQYAOFTLEKIKENYHVQVQOOH 513

RESULT 2

Q66661 PRELIMINARY; PRT; 414 AA.
AC Q66661:
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)

DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE DNA POLYMERASE PROCESSIVITY SUBUNIT.
 OS Equine herpesvirus 2.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Betaherpesvirinae; Roseolovirus.
 OX NCBI_TaxID=12657;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=86/67;
 RX MEDLINE=95302501; PubMed=7783207;
 RA Telford E.A., Watson M.S., Aird H.C., Perry J., Davison A.J.;
 RT "The DNA sequence of equine herpesvirus 2";
 RL J. Mol. Biol. 249:520-528(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=86/67;
 RA Telford E.A.R.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U20824; AAC13847.1;
 SQ SEQUENCE: 414 AA; 45191 MW; BA5A691785BE441F CRC64;

Query Match 12.1%; Score 79.5; DB 14; Length 414;
 Best Local Similarity 26.5%; Pred. No. 2.8;
 Matches 22; Conservative 13; Mismatches 33; Indels 15; Gaps 3;

QY 24 ENLDPELLSAPQDVKGHCHEAFAFCQKAKLPSPGNKNTFIIDLVAQLRRRLPARRG 83
 I I : : : : I I : : : : I I : : : : I I : : : :
 Db 308 EGDVQTQEVSS---VASTCRHLSSECSLDPPRTPELPGSPDTF-----KEIPGRSG 355
 QY 84 GKQKHIAKCPSCDSYEKTPKE 106
 I : : I I : : I I I I : :
 Db 356 ----SVHLERLDCSDSEETPKQ 375

RESULT 3
 Q9SUK4
 ID Q9SUK4 PRELIMINARY; PRT; 1304 AA.
 AC Q9SUK4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE DISEASE RESISTANCE RPP5 LIKE PROTEIN.
 GN DL4490C OR AT4G16920.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 CC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terry N.,
 RA Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
 RA Puigdomenech P., Hatzipoulos P., Obermaier B., Duesterhoft A.,
 RA Jones J., Palme K., Ansorge W., Delseny M., Bancroft I., Mewes H.W.,
 RA Schueller C., Chalwatzi N.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z97342; CAB46046.1;
 DR EMBL: AL161545; CAB80963.1;
 DR INTERPRO: IPR00157;
 DR INTERPRO: IPR000767;
 DR INTERPRO: IPR001611;
 DR INTERPRO: IPR002182;
 DR PFAM: PF00560; LRR; 4.
 DR PFAM: PF00931; NB-ARC; 1.

DR PFAM: PF01582; TIR; 1.
 DR PRINTS: PRO0364; DISEASERSIST.
 SQ SEQUENCE 1304 AA; 147728 MW; BE661972EF2DD0DA CRC64;

Query Match 11.8%; Score 78; DB 10; Length 1304;
 Best Local Similarity 24.0%; Pred. No. 13;
 Matches 36; Conservative 26; Mismatches 46; Indels 42; Gaps 7;

QY 4 DRLILRLHLIDIVEQLKIYENDLDPPELL-----SAPQDVKGHCHEAFAFCQKAKL 55
 I I : : : : I I : : : : I I : : : : I I : : : :
 Db 320 DRQLLK-AHEIDLYVEVKLPQGLALQWISQYAFGKDSPPDDFKA-----LAFEVREL 371
 QY 56 KPSNP-----GNKNTFIIDLVAQLRR-----RLPARRGGKKOKHIAKCP 94
 I I : : : : I I : : : : I I : : : : I I : : : :
 Db 372 AGSLPLGLSVLGSILKGRDKDEWKMPRLRNDSDDKTEETLRVCYDRLNKKNRELFCI 431
 QY 95 SC--DSYEKTPPEFLE---RLKWLLOKMI 119
 I : : : : I I I I : : : : I I : : : :
 Db 432 ACFFNGFKVSNVKELLEDDVGLTMLVKSLL 461

RESULT 4
 Q9XWF2
 ID Q9XWF2 PRELIMINARY; PRT; 404 AA.
 AC Q9XWF2;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE Y52BIIA.9 PROTEIN.
 GN Y52BIIA.9.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 CC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Lennard N.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Willson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copey J., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
 RA Smailson N., Smith A., Sonhammer E., Staden K., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinon-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans";
 RL Nature 368:32-38(1994).
 DR EMBL: AL032654; CAA21720.1;
 DR INTERPRO: IPR000822;
 DR INTERPRO: IPR002358;
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; UNKNOWN_1.
 DR PROSITE: PS00525; RIBOSOMAL_L6_1;
 SQ SEQUENCE 404 AA; 46400 MW; 9F55A1413525F057 CRC64;

Query Match 11.7%; Score 77; DB 5; Length 404;
 Best Local Similarity 23.9%; Pred. No. 5;
 Matches 26; Conservative 19; Mismatches 34; Indels 30; Gaps 5;
 QY 26 DLDPILLSAPQDVKGHCHEAFAFCQKAKL---PSNPGNKKTFIIDLVAQLRRRLPARR 82
 I I : : I I I I : : I I : : I I : :
 Db 206 DLNLGILDKRLDKVSGVASAKISIFDMPKVKKEDPDEPGSQ-----PSRK 251

[illegible]

DR	EMBL; AE003435; AAF46065.1; -	DR	HSP; P15167; IATL	DR	FLYBASE; FBgn029791; CG4096	DR	INTERPRO; IPR000130; -	DR	INTERPRO; IPR000884; -	DR	INTERPRO; IPR001590; -	DR	INTERPRO; IPR002465; -	DR	PFAM; PF00090; tsp_1; 2	DR	PFAM; PF01421; Repolysin; 1	DR	PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1	DR	PROSITE; PS00340; RECEPTOR_CYTOKINES_2; UNKNOWN 1	SQ	SEQUENCE 1054 AA; 118616 MW; DC154555SCB6212 CRC64;
Query Match	11.4%	Score 75;	DB 5;	Length 1054;																			
Best Local Similarity	27.6%	Pred. No. 22;																					
Matches	24;	Conservative	12;	Mismatches	31;	Indels	20;	Gaps	4;														
QY	27	LDPELLSAPQDVKGHCHEAAAFACFOKALKPSNPNNKTFITDLVAQLRRRLPARRGGKK	86																				
Db	188	IEPSKEHEPVPNGH-----PHVVFQSRSSVKPKH-----SLKRNKRKRGKS	230																				
QY	87	QKHIACPCSDSYEKRTPEFLERLKW	113																				
Db	231	SGS-AEVSNCGTREPRRME--TRLEW	254																				
RESULT 10																							
Q65168	ID	Q65168	PRELIMINARY;	PRT;	354	AA.																	
AC	Q65168;																						
DT	01-NOV-1996	(TrEMBLrel. 01, Created)																					
DF	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)																					
DT	01-JUN-2000	(TrEMBLrel. 14, Last annotation update)																					
DE	P354L.																						
GN	B354L.																						
OS	African swine fever virus (ASFV).																						
OC	Viruses; dsDNA viruses, no RNA stage;																						
OC	African swine fever-like viruses.																						
OX	NCBI_TaxID=10497;																						
RN	[1]																						
RN	SEQUENCE FROM N.A.																						
RC	STRAIN=BA71V;																						
RC	MEDLINE=96036500; PubMed=7483270;																						
RA	Yanez R.J., Rodriguez J.M., Nogal M.L., Yuste L., Enriquez C.,																						
RA	Rodriguez J.F., Vinuela E.;																						
RT	"Immune protection conferred by the baculovirus-related glycoprotein																						
RT	of Thogoto virus (Orthomyxoviridae).";																						
RL	Virology 208:249-278(1995).																						
RN	[2]																						
RN	SEQUENCE FROM N.A.																						
RC	STRAIN=BA71V;																						
RC	MEDLINE=94233765; PubMed=8178480;																						
RA	La Vega I., Gonzalez A., Blasco R., Calvo V., Vinuela E.;																						

DR	INTERPRO:	IPR001789;	-
DR	PFAM:	PFO0072;	response_reg; 1.
DR	PFAM:	PF00512;	signal; 1.
DR	PRINTS:	PR00344;	BCTRLSENSOR.
SQ	SEQUENCE	1081 AA;	118904 MW; BEB0B5D9E40573B1 CRC64;

Query Match			
Best Local Similarity 11.2%; Score 74; DB 3; Length 1081;			
Matches 34; Conservative 19; Mismatches 50; Indels 20; Gaps			

QY	5	RLILRLHLIDIVEQ--LKIYE----	NLDLPILLSAPQDVKGHCHEAAAFACFOKAL--KP 57
	:	: : : : :	: : : : :
Dd	891	RILEKOGHSVEVVENGLEAYEAIKRNKYDVLMDVQMPVMG-----GFATEKIRQWEKK 945	
QY	58	SNPGNNKTFTIIDVAQLRRRLPARRGGKKOKHIAKPCSDSCSYEKTPKFELERLKWLLQK 117	
	:	: : : : :	: : : : :
Dd	946	SNPIDSLTFTPTPIALTAHAML-----GDREKSLAK--GMDDYVSKPLPKPL--LMQTINK 997	
QY	118	M1H 120	
	:	:	:
Dd	998	C1H 1000	

RESULT 14			
ID	074271	PRELIMINARY;	PRT; 1081 AA.
AC	074271;		
DT	01-NOV-1998	(TEMBLrel. 08, Created)	
DT	01-NOV-1998	(TEMBLrel. 08, Last sequence update)	
DT	01-OCT-2000	(TEMBLrel. 15, Last annotation update)	
DE	HISTIDINE KINASE.		
GN	CHIKI.		
OS	Candida albicans (Yeast).		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;		
OX	anamorphic Saccharomycetales; Candida.		
NCBI_TaxID=5476;			
[1]	SEQUENCE FROM N.A.		
RP	STRAIN=ATCC366;		
RC	MEDLINE=98284059; PubMed=9618540;		
RX	Alex L.A., Korch C., Selitrennikoff C.P., Simon M.I.;		
RT	"COSI, a two-component histidine kinase that is involved in hyphal development in the opportunistic pathogen Candida albicans.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 95:7069-7073(1998).		
RP	[2]		
RC	SEQUENCE FROM N.A.		
RP	STRAIN=ATCC366;		
RC	Agnan J., Korch C., Alex L., Simon M.I., Selitrennikoff C.P.;		
RL	Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.		
EMBL:	U69886; AAC23929.1; ..		
HSP:	P06143; IAB6.		
DR	INTERPRO: IPR000410; .		
DR	INTERPRO: IPR001789; .		
DR	PFAM: PFO0072; response_reg; 1.		
DR	PFAM: PF00512; signal; 1.		
DR	PRINTS: PR00344; BCTRLSENSOR.		
SQ	SEQUENCE 1081 AA; 118904 MW; E11D646D225D83BF CRC64;		

Query Match			
Best Local Similarity 11.2%; Score 74; DB 3; Length 1081;			
Matches 34; Conservative 19; Mismatches 50; Indels 20; Gaps			

QY	5	RLILRLHLIDIVEQ--LKIYE----	NLDLPILLSAPQDVKGHCHEAAAFACFOKAL--KP 57
	:	: : : : :	: : : : :
Dd	891	RILEKOGHSVEVVENGLEAYEAIKRNKYDVLMDVQMPVMG-----GFATEKIRQWEKK 945	
QY	58	SNPGNNKTFTIIDVAQLRRRLPARRGGKKOKHIAKPCSDSCSYEKTPKFELERLKWLLQK 117	
	:	: : : : :	: : : : :
Dd	946	SNPIDSLTFTPTPIALTAHAML-----GDREKSLAK--GMDDYVSKPLPKPL--LMQTINK 997	
QY	118	M1H 120	

Search completed: May 23, 2001, 11:49:01
Job time: 373 sec

RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA	Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA	Williams S.M., Woodgate T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA	Gibbs R.A., Myers E.N., Rubin G.M., Venter J.C.:
RT	"The genome sequence of <i>Drosophila melanogaster</i> ."
RL	Science 287:2185-2195(2000).
DR	EMBL; AE003493; AAF48314.1; --
DR	FLYBASE; FBgn0004047; Yp3.
DR	INTERPRO; IPR000734; --
DR	INTERPRO; IPR000901; --
DR	PFAM; PF00151; lipase.1.
DR	PROSITE; PS00867; CPSASE.2; UNKNOWN.1.
SQ	SEQUENCE 421 AA; 46693 MW; 5BD4AAF75B0995DE CRC64;

```

Query Match          9.0%; Score 76.5; DB 5; Length 421;
Best Local Similarity 24.4%; Pred. No. 11;
Matches 39; Conservative 14; Mismatches 54; Indels 53; Gaps 7;

QY      8  MERIVICLMVIFGLTVHKSSSQQDR-----HMIRMROLIDI----- 45
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1  MMSURICLLATCLLVAHAHASKDASNDRLKPTKWITALENVPSLNDITWERLENQPLEQ 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY     46  -----VDQLKNVYDNLVPPEFLPAPEDEVETNCWSAFSCFQKAQLK-----SA 87
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     61  GAKVIEKIYHVGQIK---HDLTPSFVSPSNVPV---WIISNGQKVECKLNNYVETAKA 114
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY     88  NTGNERNRIINVSIKKL-KRKPPSTNAGR-----QKHL 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    115  QPGFGEDVTVLTGLPKTSPAOKKAMRLLIOAVYOKYNL 154
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT	15
Q23239	
ID	PRELIMINARY; PRT; 424 AA.
AC	Q23239;
DC	01-NOV-1996 (TrEMBLrel. 01, Created)
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DD	01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE	YOLK PROTEIN 3 PRECURSOR.
OS	Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Oestroidea; Calliphoridae; Calliphora.
OX	NCBI_TaxID=7373;
RN	[1]
RX	SEQUENCE FROM N.A.
RP	MEDLINE=94275878; PubMed=8007002;
RA	Martinez A., Bownes M.;
RT	"The sequence and expression pattern of the Calliphora erythrocephala
RL	yolk protein A and B genes.";
RT	J. Mol. Evol. 38:336-351(1994).
DR	EMBL; X70795; CAA50066.1; "
DR	INTERPRO; IPR000734; "
DR	PFAM; PF00151; lipase; 1.
KW	Signal.
FT	SIGNAL 1 19
FT	CHAIN 20 424
SQ	SEQUENCE 424 AA; 46439 MW; C35ECB97131603ED CRC64;

```

Query Match      9.08; Score 76.5; DB 5; Length 424;
Best Local Similarity 24.5%; Pred. No. 11;
Matches 37; Conservative 23; Mismatches 58; Indels 33; Gaps 7;

QY 24 VHKSSSQDQRHMRMLQIDIVDLQKLVYNDLVPEFLPAPED-----VETNC 71
   | | | | | : : : : : | : : : : | : : : : | : : : : |
Db 51 LQKLESVKGAEMLQKLCDLHQIN---NDLKPSFLPSSNPPCYIVKPNGKRVSL 107

```


Matches 34; Conservative 19; Mismatches 55; Indels 45; Gaps 6

Qy 15 LMVIFGLTVHKSSSQ-----ODRHIMRMRLQIDIVDQLKNYVNDLVPFELPAPEDV 67
:
Db 52 LTIPFLCAYVPKTEANHCKWSVDLKDELIKTSEDISLYTAN-----TYEDI 100
:
Qy 68 ETNCEWSAFSCF---OKAQLKSANTGNNERIINVISIKLLKKRPPSTNAGRQKHRLT--- 121
:
Db 101 E--COEPMYRCFFLEMKVILHECD-----IKCRSKRDVRNIWKGNNARFATYQ 147
:
Qy 122 -----CPSCDSYEKKPKPEFLERFKSLLOK 146
:
Db 148 LNSTAKCKECEEEYEENFTFIQS FVKVIQR 180
: :

RESULT 12

Q9JJ48 PRELIMINARY; PRT; 305 AA.

AC Q9JJ48;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PUTATIVE ZINC FINGER PROTEIN FLIZ1.
GN FLIZ1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=LIVER;
RA Dahm K.E., Mueller A.M.;
RT "Fetal liver zinc finger 1 (Flizl), a new gene transcribed in the
fetal liver containing three repeats of a C3H zinc finger motif";
RT fetal liver containing three repeats of a C3H zinc finger motif";
RL Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF061961; AAF74513.1; -
SQ SEQUENCE 305 AA; 34922 MW; 2F45A0758D56245C CRC64;

Query Match 9.2%; Score 78; DB 11; Length 305;
Best Local Similarity 26.2%; Pred. No. 5.6;
Matches 37; Conservative 19; Mismatches 51; Indels 34; Gaps 7;

Qy 26 KSSSQG-QDRHMIRMRLQIDIVDQLKNYVNDLVPFELPAPEDVETNCESWAFSCFQAQL 84
:
Db 88 ETC SQSGSDNFDELQQYI---QAKEMANAQAQLLPPEPVKAGASGTQOTAKQN-- 141
:
Qy 85 KSANTGNNERIINVISIKLLKKRPST-----NARRQKHRLTCPSCDSYEKKRPKP 134
:
Db 142 KKSRAGHKK----VKQKMKRKPCTGDKGRALLKNSGSREQ-----TDEPEEKQPR 190
:
Qy 135 EFLERFKSLLOKMIHOHLSSR 155
:
Db 191 -----VMSQGFINQHTVER 205
: :

RESULT 13

Q9Y378 PRELIMINARY; PRT; 334 AA.

AC Q9Y378;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE CGI-68 PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Lin W.-C.;
RT "Comparative gene cloning: Identification of novel human genes with

RT	Caenorhabditis elegans proteome as template."
RL	Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF151826; AAD34063.1; -
SQ	SEQUENCE 334 AA; 38279 MW; 3AD02031AC6C83DE CRC64;
Query Match	9.2%; Score 78; DB 4; Length 334;
Best Local Similarity	24.5%; Pred. No. 6.2;
Matches	45; Conservative 26; Mismatches 63; Indels 50; Gaps 11
QY	24 VHKSSQGDHRMI-----RMRLQIDIVDLQKN-YVNDLVPEFLPA-----P 64
DB	147 LHSEDTLQMDGHILSKRYAVIGADRLDLSLEELKKCNMTQLPTLLIAECLVIMTP 206
QY	65 EDVETNCWSAFSCFQKAQL---KSANTGNERRIINVSIKKLRKPPSTNAG-----R 114
DB	207 EGSANLLKWAANS-FERAMFINVEQVNG--DRFGQIMIENT--RRPSDLAGVETCKSLE 262
QY	115 RQKHRL-----TCPSCDSYE--KKPK-----EFLERFKSLQKKMIHLLSSRTH 157
DB	263 SOKERLLSNGWETASAVDMMLYNRLPRAEVSRIESLEFLDEWELLEQLMRHYCLCWATK 322
QY	158 GSED 161
DB	323 GGNE 326
RESULT 14	
Q9VY89	PRELIMINARY; PRT; 421 AA.
ID	Q9VY89
AC	Q9VY89;
CI	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE	01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE	YP3 PROTEIN.
GN	YP3.
OS	Drosophila melanogaster (Fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
OX	NCBI_TaxID=7227;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=BERKELEY;
RX	MEDLINE=20196006; PubMed=10731132;
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Ananatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.E., Vandell M.D., Zhang Q., Chen L.X.,
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA	Artil J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA	Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Haris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA	Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA	Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA	Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RT

Df 134 QIYDSVRHTLSGLADHQVSA-VHDTDN 163

RESULT 7
Q9R2J7 PRELIMINARY; PRT; 899 AA.
ID Q9R2J7 AC Q9R2J7
DC DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE RELAXASE.
GN NIKB.
OS Shigella sonnei.
OG Plasmid Colib-p9.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Shigella.
XC NCBI_TaxId=624;
RN [1]
RP SEQUENCE FROM N.A.
RA Sampei G., Mizobuchi K.;
RT "Organization and diversification of plasmid genomes: complete
nucleotide sequence of the Colib-p9 genome.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB021078; BAA75140.1; - .
KW Plasmid.
SQ SEQUENCE 899 AA; 103977 MW; E7DB0164C54914E3 CRC64;

Query Match 9.8%; Score 83; DB 2; Length 899;
Best Local Similarity 22.5%; Pred. No. 5.8;
Matches 34; Conservative 22; Mismatches 55; Indels 40; Gaps

Qy 27 SSSGQRHMIRMLQLDIDVDLKNYVNDLVPEFLPAPEDVTNCESAF---SCFOKAQ 83
||||| : | : | : | : | : | : | : | : | : ||| : | : | :
Db 38 SSSAQEPHSRSRLVDYATRLN-----ESFVALDVNMKGCEWVFVGTCFHCT 91
||| : | : | : | : | : | : | : | : | : | : | : | : | : | :

Qy 84 LKSANTGNRIINVSITKLKRPPSTNAGRGRKHL---TCP-----SCSYEKKKPK 134
||| : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 92 SLETAADMVEVI-----ARAHYAKDDTPVPFYILSWQSHPRE 133

Qy 135 EFLEFRFSLLQKM---IHQILLSRHSGSDS 162
: : : : : | : | : | : | : | : | : | : | : | : | :
Db 134 QIYDSVRHTLSGLADHQVSA-VHDTDN 163

RESULT 8
Q9YE8 PRELIMINARY; PRT; 566 AA.
ID Q9YE8 AC Q9YE8
DC DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG15745 PROTEIN.
GN CG15745.

OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Phrygana; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
XC NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BERKELEY,
RC MEDLINE=20196006; PubMed=107311132;
RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantades P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champé M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglou L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhargava D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

1	99.5	11.7	114	4	Q9UBA3	Q9UBA3 homo sapien
2	91.5	11.7	136	4	Q00440	Q00440 homo sapien
3	91.5	10.8	567	11	Q05208	Q05208 mus musculus
4	90.5	10.6	162	6	Q9XSJ6	Q9XSJ6 ovis aries
5	83.5	9.8	3052	14	Q82933	Q82933 johnsongrass
6	83	9.8	899	2	Q52336	Q52336 escherichia
7	83	9.8	899	2	Q9R2J7	Q9R2J7 shigella so
8	81.5	9.6	566	5	Q9VYE8	Q9VYE8 drosophila
9	81	9.5	385	5	Q9VJL7	Q9VJL7 drosophila
10	81	9.5	414	5	Q9NKG0	Q9NKG0 drosophila
11	79.5	9.4	187	13	Q9W756	Q9W756 gallus gall
12	78	9.2	305	11	Q9JJA8	Q9JJA8 mus musculus
13	78	9.2	334	4	Q9J378	Q9J378 homo sapien
14	76.5	9.0	421	5	Q9VY89	Q9VY89 drosophila
15	76.5	9.0	424	5	Q23239	Q23239 calliphora
16	75.5	8.9	336	11	Q62612	Q62612 rattus norv
17	75.5	8.9	406	5	Q22975	Q22975 caenorhabdi
18	75.5	8.9	566	11	Q62611	Q62611 rattus norv
19	75	8.8	321	4	Q9NTG4	Q9NTG4 homo sapien


```

CC -!- PATHWAY: THIRD STEP IN THE DEGRADATION OF ALLANTOIN (PURINE
CC CATABOLISM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M64778; AAA73025.1; -
DR EMBL; Z38061; CAA86192.1; -
DR PIR; S42022; S42022.
DR PIR; S48494; S48494.
DR SGD; S0001471; DAL3.
KW Hydrolase; Purine metabolism; Prenylation; Lipoprotein.
FT LIPID 192 192
SQ SEQUENCE 195 AA; 21727 MW; 6230AEE69585206B CRC64;

Query Match 8.9%; Score 76; DB 1; Length 195;
Best Local Similarity 27.6%; Pred. No. 3;
Matches 24; Conservative 14; Mismatches 29; Indels 20; Gaps 3;

QY 24 VKSSSGQDRHMRQLIDIVDLQKYNVNDLVPEFLPAPEDVETNCWSAFSCFQKQAQ 83
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B 33 LEKGANQGT---AIKLQ-----VSQVENKTSKVP-----NNWLFRCFPQPH 72
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 84 LKSANTGNNERIINVSIIKKLRKPPST 110
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 73 LNRVFTQGSNQAIASHSIKVLKHPKST 99
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
IL4_FELCA STANDARD; PRT; 133 AA.
AC P5030; P79170; O62774;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-4 PRECURSOR (IL-4) (B-CELL STIMULATORY FACTOR 1) (BSF-1)
DE (LYMPHOCYTE STIMULATORY FACTOR 1).
GN IL4.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral blood;
RA Schjins V.E.C.J., Wierda C.M.H., van Dam E.J.M., Vahlenkamp T.W.,
RA Horzinek M.C.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Lerner D.L., Elder J.H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 25-126 FROM N.A.
RA Harley R., Helps C.R., Gruffydd-Jones T.J., Day M.J., Harbours D.A.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: IL-4 PARTICIPATES IN AT LEAST SEVERAL B-CELL ACTIVATION
CC PROCESSES AS WELL AS OF OTHER CELL TYPES. IT IS A COSTIMULATOR OF
CC DNA-SYNTHESIS. IT INDUCES THE EXPRESSION OF CLASS II MHC MOLECULES
CC ON RESTING B-CELLS. IT ENHANCES BOTH SECRETION AND CELL SURFACE
CC EXPRESSION OF IGE AND ICG1. IT ALSO REGULATES THE EXPRESSION OF
CC THE LOW AFFINITY FC RECEPTOR FOR IGE (CD23) ON BOTH LYMPHOCYTES
CC AND MONOCYTES (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IL-4 / IL-13 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X87408; CAA60856.1; -
DR EMBL; U39634; AAB42052.1; -
DR EMBL; AF054602; AAC15975.1; -
DR HSP; P05112; ICYL.
DR InterPro; IPR001325; -
DR InterPro; IPR002354; -
DR Pfam; PF00727; IL4; 1.
DR PRINTS; PR00431; INTERLEUKIN4.
DR PROSITE; PS00838; INTERLEUKIN_4_13; 1.
KW Cytokine; Glycoprotein; B-cell activation; Growth factor; Signal.
FT SIGNAL 1 24
FT CHAIN 25 133
FT DISULFID 48 85
FT DISULFID 70 113
FT CARBOHYD 28 28
FT CARBOHYD 45 45
FT CARBOHYD 62 62
FT CARBOHYD 84 84
FT CARBOHYD 96 96
FT CARBOHYD 102 102
FT CONFLICT 2 2
FT CONFLICT 5 5
SQ SEQUENCE 133 AA; 15148 MW; 6D94C71F61EFC75 CRC64;

Query Match 8.9%; Score 75.5; DB 1; Length 133;
Best Local Similarity 26.6%; Pred. No. 2.1;
Matches 38; Conservative 25; Mismatches 57; Indels 23; Gaps 8;

QY 12 VICLMVIFGLTVLHKSSGQDRHMRQLIDIVDLQKYNVNDLVPEFLPAPEDVETN 70
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B 11 LVCLLA-FTSTFVHGQNFNTLKEIK---TLNLTARNDSMELTVMDVLAAPK---N 62
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 71 CEWSAFSCFQKQAQLKASANTGNNERIINVSIIKKL--RPPSTNAGRQRKHRTCPSCDSY 128
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B 63 TSDKEIFCRATTVLRQIYTH-----NCSTKFLGLDRNLSSMANR-----TC-SVNEV 110
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 129 EKKPKPEFLERFKSLLOKMIHOH 151
D : : : : : : : : : : : : : : : : : : : : : : : : : : : :
B 111 KCTLKDFLERLKAIMQKYSKH 133
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
FM14_MOUSE STANDARD; PRT; 1206 AA.
AC Q05859;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE FORMIN 1 ISOFORM IV (LIMB DEFORMITY PROTEIN).
GN FMN OR LD.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA MEDLINE=92112033; PubMed=1339380;
RA Grubby-Jackson L., Kuo A., Leder P.;
FT "A variant limb deformity transcript expressed in the embryonic mouse
FT limb defines a novel formin.";
RL Genes Dev. 6:29-37(1992).
CC -!- FUNCTION: IS IMPORTANT IN THE MORPHOGENESIS OF LIMB AND MAY HAVE A
CC FUNCTION IN DIFFERENTIATED CELLS OR BE INVOLVED IN MAINTAINING
CC SPECIFIC DIFFERENTIATED STATES.

```



```

DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOHETICAL 23.8 KDA PROTEIN IN MET18-STH1 INTERGENIC REGION.
GN YIL127C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Barrell B.G., Badcock K., Bankier A.T., Bowmen S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gents S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z46833; CAA86865.1; -
DR SGD; S0001389; YIL127C.
KW Hypothetical protein.
SQ SEQUENCE 206 AA; 23845 MW; 4E51B11216F7ECB8 CRC64;

Query Match          9.3%; Score 79; DB 1; Length 206;
Best Local Similarity 26.1%; Pred. No. 1.7;
Matches 29; Conservative 18; Mismatches 48; Indels 16; Gaps 3;

QY 48 QLNKYNVDLPVPEFLPAEDVETNCWAFSCFQAKLSANTGNNRIINYSIKLKRKP 107
DB 12 QATSVVNGLSNLLPGVPKTRANNKTSVNGSKAQLIDRLNLRVQLQNDVHKIKKC 71

QY 108 PSTNAGRRQKRLHRTCPSCDSYEKKPPKEFLERF--KSLQKMIHQHLSSRT 156
DB 72 KLVKKKKVKKHKL-----DREQLQAKHQVLLKK--HQEGTTLT 108

RESULT 10
IDH_AZOV1
ID IDH_AZOV1 STANDARD; PRT; 741 AA.
AC P16100;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE ISOCITRATE DEHYDROGENASE [NADP] (EC 1.1.1.42) (OXALOSUCCINATE
DE DECARBOXYLASE) (IDH).
GN ICD.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RA Sahara T., Suzuki M., Tsuruha J.I., Takada Y., Abe K., Fukunaga N.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP PRELIMINARY SEQUENCE OF 228-250 AND 254-259.
RX MEDLINE=74086945; PubMed=4149369;
RA Edwards D.J., Heinrikson R.L., Chung A.E.;
RT "Triphosphopyridine nucleotide specific isocitrate dehydrogenase from
RT Azotobacter vinelandii. Alkylation of a specific methionine residue
RT and amino acid sequence of the peptide containing this residue.";
RL Biochemistry 13:677-683(1974).
CC -|- CATALYTIC ACTIVITY: ISOCITRATE + NADP(+) -> 2-OXOGLUTARATE +

```

```

CC CO(2) + NADPH.
CC -|- ENZYME REGULATION: INHIBITION OF THIS ENZYME BY PHOSPHORYLATION
CC REGULATES THE BRANCH POINT BETWEEN THE KREBS CYCLE AND THE
CC GLYOXYLATE BYPASS, WHICH IS AN ALTERNATE ROUTE THAT ACCUMULATES
CC CARBON FOR BIOSYNTHESIS WHEN ACETATE IS THE SOLE CARBON SOURCE
CC FOR GROWTH. THE PHOSPHORYLATION STATE OF THIS ENZYME IS CONTROLLED
CC BY ISOCITRATE DEHYDROGENASE KINASE/PHOSPHATASE (ACEK).
CC -|- SUBUNIT: MONOMER.
CC -|- SIMILARITY: BELONGS TO THE MONOMERIC-TYPE FAMILY OF IDH.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D73443; BA011169.1; -
DR PIR; A10759; A10759.
KW Oxidoreductase; NADP; Phosphorylation; Glyoxylate bypass;
KW Tricarboxylic acid cycle.
SQ SEQUENCE 741 AA; 80389 MW; 29FF35278B5AED8B CRC64;

Query Match          9.2%; Score 78; DB 1; Length 741;
Best Local Similarity 24.8%; Pred. No. 8.9;
Matches 26; Conservative 23; Mismatches 34; Indels 22; Gaps 4;

QY 33 DRHMIRMRLQIDIVDLQKLNKYNVDL-----VPEFLPAEDVETNCWAFSCFQAKLS 86
DB 77 DANIILKLPNISASVQLKAAIKELQOQGYKLPDY---PEEPKTDTEKDYKARYDKIGSA 133

QY 87 AN-----TGNNRIINYSIKLKRKPSTNAGRRQKRLHRTCPSCDS 127
DB 134 VNPVREGNSRRAPLSVKNYARKHP-----HKMGAWGADS 169

RESULT 11
VIT3_DROME
ID VIT3_DROME STANDARD; PRT; 420 AA.
AC P06607;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE VITELLOGENIN III PRECURSOR (YOLK PROTEIN 3).
GN YP3.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RX MEDLINE=87305580; PubMed=3114046;
RA Garabedian M.J., Shlirra A.D., Bownes M., Wensink P.C.;
RT "The nucleotide sequence of the gene coding for Drosophila
RT melanogaster yolk protein 3.";
RL Gene 55:1-8(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87146365; PubMed=3029679;
RA Yan Y.L., Kunert C.J., Postlethwait J.H.;
RT "Sequence homologies among the three yolk polypeptide (Yp) genes in
RT Drosophila melanogaster.";
RL Nucleic Acids Res. 15:67-85(1987).
RN [3]
RP SEQUENCE FROM N.A. (MUTANT YP3S1).
RX MEDLINE=91360094; PubMed=1909425;
RA Liddell S., Bownes M.;
RT "Characterization, molecular cloning and sequencing of YP3s1, a
RT fertile yolk protein 3 mutant in Drosophila.";

```


use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL; Y07519; CAA68812.1; -.
CC EMBL; X60184; CAA42742.1; -.
CC EMBL; M24843; AAA40160.1; -.
CC PIR; S07054; S07054.
CC MGD; MGI:98427; Ly84.
CC InterPro; IPR003006; -.
CC Pfam; PF00047; 19; 3.
KW Immunoglobulin domain; Glycoprotein; Signal.
FT SIGNAL 1 26
FT CHAIN 27 337
FT DOMAIN 35 100
FT DOMAIN 132 194
FT DOMAIN 233 315
FT DISULFID 42 93
FT DISULFID 139 187
FT DISULFID 240 308
FT CARBOHYD 60 60
FT CARBOHYD 101 101
FT CARBOHYD 107 107
FT CARBOHYD 146 146
FT CARBOHYD 176 176
FT CARBOHYD 194 194
FT CARBOHYD 225 225
FT CARBOHYD 259 259
FT CARBOHYD 278 278
FT VARIANT 192 192
SQ SEQUENCE 337 AA; 38502 MW; 7574372722486926 CRC64;

```

Query Match 10.8%; Score 91.5; DB 1; Length 337;
 Best Local Similarity 25.6%; Pred. No. 0.21;
 Matches 33; Conservative 20; Mismatches 39; Indels 37; Gaps 7;

```

QY 47 DQKKNVNDLVPEFLPAPEDVETNCWFAFCFQKALQSANTGNINIVSIKKLKRK 106
DB 76 DRLK-----FLPAPVE-----DSGIYACV-----IRSPNL-NKTGYLVNTIHK---K 113
QY 107 PPSTN-----AGRRQKRLTFCPSDCSYEKPKPKFLEFRKSLQKMIHQH----- 151
DB 114 PPSNIPDLYMSTVGRGSKFKITCTIDLYNTAPVQWFKNCKALQEPFRHRSYLF 173
QY 152 LSSRTHGSE 160
DB 174 IDNVTHDDE 182

```

RESULT 6
 IL15_BOVIN
 ID IL15_BOVIN STANDARD; PRT; 162 AA.
 AC Q28028;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE INTERLEUKIN-15 PRECURSOR (IL-15).
 GN IL15.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HOLSTEIN;
 RX MEDLINE=97426124; PubMed=9282828;
 RA Canals A., Gasbarre L.C., Boyd P.C., Almeria S., Zarlenga D.S.;
 RT "Cloning and expression of bovine interleukin-15: analysis and
 RT modulation of transcription by exogenous stimulation.";
 RL J. Interferon Cytokine Res. 17:473-480(1997).

-!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15 WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R GAMMA BUT NOT IL-2R ALPHA (BY SIMILARITY).
 -!- SUBCELLULAR LOCATION: SECRETED.
 -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL; U42433; AAA85130.1; -.
CC Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29
FT PROPEP 30 48
FT CHAIN 49 162
FT INTERLEUKIN-15.
FT DISULFID 83 133
FT DISULFID 90 136
FT CARBOHYD 104 104
FT CARBOHYD 121 121
FT CARBOHYD 121 121
FT CARBOHYD 121 121
FT CARBOHYD 127 127
FT CARBOHYD 127 127
SQ SEQUENCE 162 AA; 18554 MW; 6A63CAA329EB8302 CRC64;

```

Query Match 10.6%; Score 90.5; DB 1; Length 162;
 Best Local Similarity 25.5%; Pred. No. 0.11;
 Matches 28; Conservative 17; Mismatches 58; Indels 7; Gaps 3;

```

QY 45 IVDQKKNVNDLVPEF-----LPAPEDVETNCWFAFCFQKALQSANTGNINIVS 99
DB 53 VINDLKT-IEHLTQSIHMDATLYTESDAHFNCKVTAMQCFLELRVLHESKNATYIEI- 110
QY 100 IKKLRKPPSTNAGRRQKRLTFCPSDCSYEKPKPKFLEFRKSLQKMIHQH 149
DB 111 IENLTMLANSNLSIENKTELGCCKECEELEKSIKEFLKSFVHIVOMFIN 160

```

RESULT 7
 ALB2_XENLA
 ID ALB2_XENLA STANDARD; PRT; 607 AA.
 AC P14872;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE 74 KDA SERUM ALBUMIN PRECURSOR.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE OF 3-607 FROM N.A.
 RX MEDLINE=89313788; PubMed=2747653;
 RA Moskatis J.E., Sargent T.D., Smith L.H. Jr., Pastori R.L.,
 RA Schoenberg D.R.;
 RT "Xenopus laevis serum albumin: sequence of the complementary
 RT deoxyribonucleic acids encoding the 68- and 74-kilodalton peptides
 RT and the regulation of albumin gene expression by thyroid hormone
 RT during development.";
 RL Mol. Endocrinol. 3:464-473(1989).
 RN [2]
 RP SEQUENCE OF 1-48 FROM N.A.
 RX MEDLINE=88172470; PubMed=2451026;
 RA Schorpp M., Doebbeling U., Wagner U., Ryffel G.U.;
 RT "5'-flanking and 5'-proximal exon regions of the two Xenopus albumin
 RT genes. Deletion analysis of constitutive promoter function.";
 RL J. Mol. Biol. 199:83-93(1988).
 RN [3]

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: U03099; AAA18416.1; -
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 162 AA; 18222 MW; 1BF9A82644E1C9B7 CRC64;

Query Match 11.1%; Score 94.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.047;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 43 IDIVDLQKNYVNDLVPF-----LPAPDEVETNCWSAFSCF-----QKAQLKSANTGNNE 93

DB 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVISHESGDTDHD 109

QY 94 RIINVSIKKLRKPPSTNAGRQKRLTCPCSDSYEKKPKPEFLERFKSLQKMIH 149

DB 110 TVENLII--LANNILSSNGNITES---GCKECELEEKNIKEFLQSFVHVQMFN 160

RESULT 4

IL15_MACMU
ID IL15_MACMU STANDARD; PRT; 162 AA.

AC P48092;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE INTERLEUKIN-15 PRECURSOR (IL-15).

GN IL15.

OS Macaca mulatta (Rhesus macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecinae; Macaca.

OX NCBI_TaxID=9544;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Blood;

RX MEDLINE=96003435; PubMed=7561102;

RA Villinger F.J., Brar S.S., Mayne A.E., Chikkala N., Ansari A.A.;

RT "Comparative sequence analysis of cytokine genes from human and

nonhuman primates.";

RL J. Immunol. 155:3946-3954(1995).

RN [2]

RP SEQUENCE FROM N.A.

RA Tatsumi M.;

RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-

LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15

WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R

GAMMA BUT NOT IL-2R ALPHA.

CC -1- SUBCELLULAR LOCATION: SECRETED.

CC -1- SIMILARITY: BELONGS TO THE IL-15 FAMILY.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: U19843; AAB60398.1; -

CC EMBL: AB000555; BAA19149.1; -

KW Cytokine; Glycoprotein; Signal.

FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48 POTENTIAL.
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 5 5 K -> T.
FT VARIANT 31 31 I -> T.
SQ SEQUENCE 162 AA; 18194 MW; D233CF7FF6188C01 CRC64;

Query Match 11.1%; Score 94.5; DB 1; Length 162;

Best Local Similarity 25.9%; Pred. No. 0.047;

Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 43 IDIVDLQKNYVNDLVPF-----LPAPDEVETNCWSAFSCF-----QKAQLKSANTGNNE 93

DB 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVISHESGDTDHD 109

QY 94 RIINVSIKKLRKPPSTNAGRQKRLTCPCSDSYEKKPKPEFLERFKSLQKMIH 149

DB 110 TVENLII--LANNILSSNGNITES---GCKECELEEKNIKEFLQSFVHVQMFN 160

RESULT 5

IRL1_MOUSE

ID IRL1_MOUSE STANDARD; PRT; 337 AA.

AC P14719;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE INTERLEUKIN 1 RECEPTOR-LIKE 1 PRECURSOR (ST2 PROTEIN) (T1 PROTEIN)

DE (LYMPHOCYTE ANTIGEN 84).

GN IL1RL1 OR ST2 OR ST2 OR LY84.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C;

RX MEDLINE=90092495; PubMed=2532153;

RA Tomimaga S.;

RT "A putative protein of a growth specific CDNA from BALB/c-3T3 cells

is highly similar to the extracellular portion of mouse interleukin 1

receptor.";

RL FEBS Lett. 258:301-304(1989).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C3H/HE; TISSUE=Spleen;

RX MEDLINE=91355215; PubMed=1832015;

RA Tomimaga S.I., Jenkins N.A., Gilbert D.J., Copeland N.G.;

RT Tetsuka T.;

RT "Molecular cloning of the murine ST2 gene. Characterization and

chromosomal mapping.";

RL Biochim. Biophys. Acta 1090:1-8(1991).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=89345536; PubMed=2527364;

RA Klemenz R., Hoffmann S., Werenskiold A.K.;

RT "Serum- and oncoprotein-mediated induction of a gene with sequence

similarity to the gene encoding carcinoembryonic antigen.";

RL Proc. Natl. Acad. Sci. U.S.A. 86:5708-5712(1989).

CC -1- FUNCTION: POSSIBLY INVOLVED IN REGULATION OF T-LYMPHOCYTE

ACTION.

CC -1- DEVELOPMENTAL STAGE: GROWTH-SPECIFIC EXPRESSION, G1-PHASE OF

CELL CYCLE.

CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC -1- SIMILARITY: STRONG, TO INTERLEUKIN-1 RECEPTORS.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2001, 09:38:18 ; Search time 9.55 Seconds
(without alignments)
581.088 Million cell updates/sec

Title: US-09-522-217-2

Perfect score: 850

Sequence: 1 MRSSPGNMERIVICLVFL.....LLQKMIHQHLSRTHGSEDS 162

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	DB ID	Description
1	106.5	12.5	1 IL15_FELCA	O97687 felis silve
2	99.5	11.7	1 IL15_HUMAN	P40933 homo sapien
3	94.5	11.1	1 IL15_CERAE	P40221 cercopithe
4	94.5	11.1	1 IL15_MACMU	P48092 macaca mula
5	91.5	10.8	1 IRL1_MOUSE	P14719 mus musculu
6	90.5	10.6	1 IL15_BOVIN	Q28028 bos taurus
7	82	9.6	1 ALB2_XENLA	PI4872 xenopus lae
8	80.5	9.5	1 IL15_PIG	Q95253 sus scrofa
9	79	9.3	1 YIM7_YEAST	P40470 saccharomyc
10	78	9.2	1 IDH_AZOVI	P16100 azotobacter
11	76.5	9.0	1 VIT3_DROME	P06607 drosophila
12	76.5	9.0	1 YG3M_YEAST	P48237 saccharomyc
13	76	8.9	1 IL15_YEAST	P32459 saccharomyc
14	75.5	8.9	1 IL15_FELCA	P50303 felis silve
15	75.5	8.9	1 FM14_MOUSE	Q05859 mus musculu
16	75	8.8	1 LCMT_HUMAN	Q9uic8 homo sapien
17	74.5	8.8	1 FUSE_ARATH	P45432 arabidopsis
18	74.5	8.8	1 YK25_CAEEL	P34332 caenorhabdi
19	73.5	8.6	1 IL15_MOUSE	P48346 mus musculu
20	73.5	8.6	1 IFT2_HUMAN	P09913 homo sapien
21	73.5	8.6	1 CCB3_RABIT	P54286 oryctolagus
22	73.5	8.6	1 CCB3_HUMAN	P54284 homo sapien
23	73.5	8.6	1 CCB3_RAT	P34287 rattus norv
24	73	8.6	1 IF2A_YEAST	P20459 saccharomyc
25	72.5	8.5	1 IL15_RAT	P97604 rattus norv
26	72	8.5	1 Y069_NPVAC	P41469 autographa
27	71.5	8.4	1 CCB3_MOUSE	P54285 mus musculu
28	71	8.4	1 AHR_MOUSE	P30561 mus musculu
29	70	8.2	1 IFT3_MOUSE	Q64345 mus musculu
30	70	8.2	1 BAC1_MOUSE	P97302 mus musculu
31	70	8.2	1 YMP3_CAEEL	Q10947 caenorhabdi
32	70	8.2	1 SMC3_YEAST	P47037 saccharomyc
33	70	8.2	1 BUD3_YEAST	P25558 saccharomyc

34	69.5	8.2	219	1	Y413_RICPR	Q9zdb9 rickettsia
35	69.5	8.2	1468	1	FN1_MOUSE	Q05860 mus musculu
36	69.5	8.2	1750	1	Y832_METJA	Q58242 methanococc
37	69	8.1	132	1	IL4_CANFA	Q77762 canis fami
38	69	8.1	568	1	PRP3_BUCAP	O85297 buchnera ap
39	69	8.1	932	1	PMS1_HUMAN	P54277 homo sapien
40	68.5	8.1	464	1	N2B_HAEIR	P46441 haematobia
41	68.5	8.1	655	1	YDH2_SCHPO	Q92347 schizosacch
42	68.5	8.1	2749	1	IP3R_RAT	P29994 rattus norv
43	68	8.0	201	1	SPC2_HUMAN	P34741 homo sapien
44	68	8.0	573	1	PTL_BORBU	O51508 borrelia bu
45	68	8.0	1172	1	PHYB_ARATH	P14713 arabidopsis

ALIGNMENTS

RESULT 1

IL15_FELCA

AC IL15_FELCA STANDARD; PRT: 162 AA.

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE INTERLEUKIN-15 PRECURSOR (IL-15).

GN IL15.

OS Felis silvestris catus (Cat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

OX NCBI_TaxID=9685;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Lymph node;

RA Barger A.B., Dean G.A., Lavoy A.S.;

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

CC FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-

CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15

CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R

CC GAMMA BUT NOT IL-2R ALPHA (BY SIMILARITY).

CC SUBCELLULAR LOCATION: SECRETED.

CC SIMILARITY: BELONGS TO THE IL-15 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF108148; AAD05268.1;

KW Cytokine; Glycoprotein; Signal.

FT SIGNAL 1 29 POTENTIAL.

FT PROPEP 30 48 POTENTIAL.

FT CHAIN 49 162 INTERLEUKIN-15.

FT DISULFD 83 133 POTENTIAL.

FT DISULFD 90 136 POTENTIAL.

FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 162 AA; 18412 MW; D8C7CE7F40110DD CRC64;

Query Match 12.5%; Score 106.5; DB 1; Length 162;

Best Local Similarity 26.0%; Pred. No. 0.0036;

Matches 38; Conservative 25; Mismatches 58; Indels 25; Gaps 6;

QY 14 CLAWFLGLTGL---VHKSSSQGQDRHMIRKQLDIVDQLKNYVNDLVPELPAPEDVETN 70

DB 30 CIPVFTILSCINAGLPKTEANWQD--VISDKIIDIKTIQSLHIDATLYTE-----SDVHFN 82

QY 71 CEWSAFSCF---QRAQLKSANTGNNERINIVSIKKLRKPPSTNAGRQKRLT---CP 123

DB 83 CKVTAMKCELLLEHLVLSKSKNETIHQTVENII-----LANSGLSSNRNRTETGCK 134

```

Db 377 ITAGTSGHSSIS-----DYNKELAEPMNEV---FAPTHPQLIQTKA----- 414
QY 77 SCFQKAQLKSANTGNNERIINV-----SIKKLK-RKPPSTNAGRQKHRLTCPSCD 126
Db 415 SAMNSNAVSPANTNNSDEIECVYGGPKTPASIKRLQPTGPRTLSCGORMKSILSVPAHQ 474
QY 127 SYEKKPPKEFLERFKSLQKMIHQ 150
Db 475 RRSIPPVKTEDEAMECLSMMYE 498

```

Search completed: May 24, 2001, 09:38:15
Job time: 37 sec


```

Qy      115  RQKHLRTPCSDSYEKKPKPEFLERFKSLQKMIHQHLSSRTHGSEDS 162
      - - - - - 1 - - - - - 1 - - - - - 1 - - - - - 1 - - - - -
Db      1068 EQQMKLVC-----KSPREYLOPFKDKLEEFFKK--AKKEHKMEES 1106

```

RESULT 12

146352
hypothetical protein DKFZp434C0816.1 - human (fragment)

cytochemical protein expression in human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
 C:Accession: T46352
 R:Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z23037

A: Accession: T46352
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-321 <AAA>
A: Cross-references: EMBL:AL137283
A: Experimental source: adult testis; clone DKFZp434C0816
C: Genetics:
A: Note: DKFZp434C0816.1

Query Match	8.8%	Score 75;	DB 2;	Length 321;
Best Local Similarity	23.0%;	Pred. No. 16;		
Matches 42;	Conservative	29;	Mismatches	64;
Indels	48;	Gaps	10;	

QY 24 VHKSQQGDRHMT-----RMQLIDIVDLQKN-YVNDLVPEFLPA-----P 64
:
: : : : : : : : : : : : :
dD 134 LHS EDTLMDGHILDSKRYANIGADIRLSELEFKKKCNMNTOTPTLTAECVIYVMTP 193

```
65 EDVETNCESAFSCFOKAQL---KSANTGNERNRIINVSIKKLKRKP-----PSTNAGR 115
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
194 EQSANKLWKAANS-FRAMEFNTYEOVMNG--DRFGQIMENLRROCDLAGVETCKSLSS 250
```

```

QY 116 QKURL-----TCPSCDSYE--KKPKP-----EFLERFKSLQKMIHQHLSRTHG 158
    |||      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
Db 251 QKERLLSNGWETASADMMELYNRLPRAEVSRYTESLEFDJDEMLLBOLMRHYCLCWATKG 310

```

QY	159	SED	161
		:	
Db	311	GNE	313

```

RESULT 13 .
F82881
hypothetical protein UU505 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: F82881
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Hainer, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A:Reference number: A82870
A:Accession: F82881
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-286 >GLA>
A:Cross-references: GB:AE002149; GB:AF222894; NID:g6899503; PIDN:AAF30917.1; GSPDB:GN001
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: UU505
A:Genetic code: SGC3

```

Query Match 8.8%; Score 74.5; DB 2; Length 286;
Best Local Similarity 20.6%; Pred. No. 16;
Matches 42; Conservative 38; Mismatches 59; Indels 65; Gaps 9;

```
QY      6 GNNERIVICLMVIFLGLTVHKS-----SSQGQRHMIRMQLIDIVD 47
```

DB ZI GGIISLLVITATILGVILSTQNTNRISSKREPLTQIEDNINNKRQDAH--KVNENITIID 78

Qy	48	QLK-----	---NVVN---	DLVPBFLPAPEDVETNC	EWSAFSCFQKAQLKSANTGN	92
		:	:	:	:	
		:	:	:	:	
		:	:	:	:	
Db	79	KKGSIHKKTIDTKINQ	KEILQKYLKLNKEE	-----	NKDQINNTNENLN	126
		:	:	:	:	
Qy	93	ERIINV-----	---SIKKLRKPPSTNAGR	KOHLKTCPCDSCDSEYKPK	PEFLERFKS	142
		:	:	:	:	
		:	:	:	:	
		:	:	:	:	
Db	127	KPIINVNVDDKKNQEN	STKLKNNDFISNDKNNK	---INENNISYEKPPK	--LKRFINI	182
		:	:	:	:	
Qy	143	L-----	LOKMHIHLSRTHGSEDS			162
		:	:	:	:	
		:	:	:	:	
Db	183	INVIANKNDYQLLDFF	KTYFESN			206

RESULT 14

S41211
voltage-dependent calcium channel protein - human
C:Species: Homo sapiens (man)
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1999
C:Accession: S41211
R:Collin, T.; Lory, P.; Taviaux, S.; Courtieu, C.
Eur. J. Biochem. 220, 257-262, 1994
A:Title: Cloning, chromosomal location and function
of a human voltage-dependent calcium channel protein.
A:Reference number: S41211; MUID:94164167

Query Match 8.8%; Score 74.5; DB 2; Length 482;
Best Local Similarity 20.8%; Pred. No. 28;
Matches 36: Conservative 27; Mismatches 65; Indels 4

```
Qy 9 ERIVICLWVIFGLTVHKSSSQGDRIHMRROLIDVDQLKN--YVNDLVPE-----FL 61
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 59 EFAVRITNVSYGVLDEECPVQGSVN-FEAKDFLHIKEYSNWDWIRGLVKEGGDAFI 117
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 62 PAPEDVTNCWSAFSCFQAQKASANTGNERRIINVISIKLRKPSTNAGRQRHR-- 119
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 118 PSPORLE-----SIRLKOEQARRSGNSPSSLDIG---NRRSPPSLAKQKQAAEH 166
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 120 -----LTCPSDSEYEKKPKPEELERFKLSLOKMTHQHLLSSHTG 158
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 167 VPPDYDVVPSMRPVVLGVPSLKGYS-----VTDMMOKALFDLFKLHRFDG 209
```

RESULT 15

S44759
C14B9.5 protein - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 14-Sep-1994 #sequence_revision 12-May-1995 #tex
C:Accession: S44759
R:Favello, A.D.
submitted to the EMBL Data Library, May 1993
A:Description: Sequence of the *C. elegans* cosmid C14B9.
A:Reference number: S44617
A:Accession: S44759
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-789 <FAV>
A:Cross-references: EMBL:L15188; NID:g289640; PID:g28966
C:Genetics:
A:Introns: 61/3; 129/2; 147/3; 191/3; 279/3; 368/3; 392/

Query Match 8.8%; Score 74.5; DB 2; Length 789;
Best Local Similarity 24.3%; Pred. No. 47;

QY 18 IFGLTVHKSSSGODRHMIRMROLIDIVDOLKNYVNDLVPEFLPA - PEDVETNCEWSAF 76

C;Species: Saccharomyces cerevisiae
C;Date: 31-Mar-1992 #sequence_revision 14-Sep-1994 #text_change 21-Jul-2000
C;Accession: S42022; S48494
R;Yoo, H.S.; Cooper, T.G.
Yeast 7, 693-698, 1991
A;Title: The ureidolysate hydrolase (DAL3) gene in Saccharomyces cerevisiae.
A;Reference number: S42022; MUID:9213160
A;Accession: S42022
A;Molecule type: DNA
A;Residues: 1-195 <YOO>
A;Cross-references: EMBL:M64778; NID:q171369; PIDN:AAA73025.1; PID:q171370
R;Rowley, K.
Submitted to the EMBL Data Library, October 1994
A;Reference number: S48478
A;Accession: S48494
A;Molecule type: DNA
A;Residues: 1-195 <ROW>
A;Cross-references: GB:247047; EMBL:Z38061; NID:g603997; PID:g763377; MIPS:YIR032C
C;Genetics:
A;Gene: SGD:DAL3
A;Cross-references: SGD:S0001471; MIPS:YIR032C
A;Map position: 9R
C;Keywords: allantoin degradation; hydrolase; lipoprotein; methylated carboxyl end; pred
F;192/Binding site: farnesyl (Cys) (covalent) #status predicted
F;192/Modified site: methyl ester carboxyl end (Cys) (in mature form) #status predicted

Query Match 8.9%; Score 76; DB 2; Length 195;
Best Local Similarity 27.6%; Pred. No. 7.7;
Matches 24; Conservative 14; Mismatches 29; Indels 20; Gaps 3;
QY 24 VHKSSGQDRHIRMRLQIDIVDLKYNVNDLPELPAPEDVETNCESAFSCFQKQAQ 83
Db 33 LKGANQGT---AIKLLQ-----VSQVENKTSKVP-----NWNLFRCFPQPH 72

QY 84 LKSANTGNNEIRINVSITKLRKPPST 110
Db 73 LNRVFTGSGNQAIASHIKVLEKHPCT 99

RESULT 9
S42632
Fit-1S protein precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 25-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 19-May-2000
C;Accession: S42632
R;Bergers, G.; Reikerstorfer, A.; Braselmann, S.; Graninger, P.; Busslinger, M.
EMBO J. 13, 1176-1188, 1994
A;Title: Alternative promoter usage of the Fos-responsive gene Fit-1 generates mRNA iso
A;Reference number: S42632; MUID:94178260
A;Accession: S42632
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-336 <BER>
A;Cross-references: GB:U04319; NID:g488278; PIDN:AAA67172.1; PID:g488279
C;Superfamily: vaccinia virus B15R protein; immunoglobulin homology

Query Match 8.9%; Score 75.5; DB 2; Length 336;
Best Local Similarity 22.2%; Pred. No. 15;
Matches 30; Conservative 26; Mismatches 42; Indels 37; Gaps 6;
QY 43 IDIVDLKYNVNDLVP-----EFLPAPEDVETNCESAFSCFQKQAQKSA 88
Db 49 INPVEWYNTNRIPIQKRNIRFVSRDRUKFLPAKVE-----DSGIYTCVIRSP-ESIK 102

QY 89 TGNNERIINVSITKLRKPP-----STNAGRQKRLTDCPSDCSDSYEKKPKPEFLE 138
Db 103 TGS----LNTVIYK---RPNCIPDYMYMISTVDGSDKSKITCPTIALYNWTAPVQWFK 155

QY 139 RFKSLQKMTQHL 153
Db 156 NCKALQGPFRRAHMS 170

RESULT 10
T28957
hypothetical protein F45F2.11 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T28957
R;Davidson, S.; Wohldmann, P.
Submitted to the EMBL Data Library, July 1996
A;Description: The sequence of C. elegans cosmid F45F2.
A;Reference number: Z20548
A;Accession: T28957
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-406 <DAV>
A;Cross-references: EMBL:U64845; PIDN:AAAC48032.1; GSPDB:GN00023; CESP:F45F2.11
A;Experimental source: strain Bristol N2; clone F45F2
C;Genetics:
A;Gene: CESP:F45F2.11
A;Map position: 5
A;Introns: 1/2; 24/2; 143/3; 307/2; 338/2

Query Match 8.9%; Score 75.5; DB 2; Length 406;
Best Local Similarity 20.5%; Pred. No. 19;
Matches 38; Conservative 23; Mismatches 55; Indels 69; Gaps 7;
QY 26 KSSSGQDRHIRMRLQIDIVDLKYNVNDLPELP----- 62
Db 13 KSSPOSSDENFKRLQWGIKRLK-----TPEFLKESSELEIEKNEKSGIKGGENEG 66
QY 63 -----APEDVETNCW-----SAFSCFQKQAQKSAKNTGNNER 94
Db 67 ENEEKCERFPVPTNEIDSHKWHYHRLMLKLEYKRGEGAPPPPPPLPSMTAA 126

QY 95 IINVSITKLR-----KPPSTNAGRQKRLTDCPSDCSDSYEKKPKPEFLERFKSLQKM 147
Db 127 VSFNAFDEVKRAAQAKTAKSPSTSSLERRAQRF-CPA-DFQPLPPPHIYIEMIRLAP-- 182

QY 148 IHQHL 152
Db 183 -HQYI 186

RESULT 11
S24407
formin isoform IV - mouse
C;Species: Mus musculus (house mouse)
C;Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C;Accession: S24407
R;Jackson-Grusby, L.; Kuo, A.; Leder, P.
Genes Dev. 6, 29-37, 1992
A;Title: A variant limb deformity transcript expressed in the embryonic mouse limb de
A;Reference number: S24407; MUID:92112033
A;Accession: S24407
A;Molecule type: mRNA
A;Residues: 1-1206 <JAC>
A;Cross-references: EMBL:X62379; NID:g51552; PIDN:CAA44244.1; PID:g51553

Query Match 8.9%; Score 75.5; DB 2; Length 1206;
Best Local Similarity 22.0%; Pred. No. 59;
Matches 37; Conservative 41; Mismatches 69; Indels 21; Gaps 8;
QY 2 RSSPG--NMERI--VICLMVIFLGLTVHKSSSQGQ-DRHMIRMRLQIDIVDLKYNV-- 54
Db 953 RASGGLLHKSVKIDILALILAFGNVNGNTRGOADGYSLIILPKLKDVKSRDNGMLV 1012

QY 55 DLVPEFTLPAPEVETNCESAFSCFQKQAQKSAKNTGNNERIINVSITKLRKPPSTNAGR 114
Db 1013 DYVVKYVLRYYDQEAGTDKSVFPLPEPDDFFLASQVFEDLLK-DLRKLKQLEAS----- 1067

Db 72 EINDFAKSCINDTPECEKPVGTLFFDKLCADPAVGNYEWSKCCAKODPERAQCFKAH 131
Qy 89 TGNERRII---NVSIKKLARK-----PPSTNAGRRQKRLTSPS 124
Db 132 RDEHTSIRKPEPETCKLKEHPDDLLSARFHEARNHPDLYPPAVLALTKQYHKLAEHC 191
Qy 125 CDSYERKPPKREFLERKSLQKMIHQHLSRTHGSED 161
Db 192 CREEDKE--KCFSEKMKQLMK-----QSHSID 217

RESULT 5
S49882
hypothetical protein Y1127c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein Y18277.02c
C:Species: Saccharomyces cerevisiae
C:Date: 13-Jan-1995 #sequence_revision 20-Feb-1995 #text_change 05-Nov-1999
C:Accession: S49882
R:Hamlyn, N.; Churcher, C.
submitted to the EMBL Data Library, November 1994
A:Reference number: S49881
A:Accession: S49882
A:Molecule type: DNA
A:Residues: 1-206 <HAM>
A:Cross-references: GB:47047; EMBL:246833; NID:g603997; PID:g763219; GSPDB:GNO0009; MIPS:Y1127c
A:Gene: MIPS:Y1127c
A:Map position: 9L

Query Match 9.3%; Score 79; DB 2; Length 206;
Best Local Similarity 26.1%; Pred. No. 4.3;
Matches 29; Conservative 18; Mismatches 48; Indels 16; Gaps 3;

Qy 48 QLKNYNDLVPEFLPAPEDVETNCWSAFSCFQKAQLKSANTGNERRIINVSIIKKLRKP 107
Db 12 QATSVVNGLLSNLLPGVPKIRANNGKTSVNVNGSKAQLIDRLNKKRVQLQNRDVHKIKKC 71
Qy 108 PSTNAGRRQKRLTSPSCDSYEKKPPKEFLERF--KSLQKMIHQHLSRST 156
Db 72 KLVRKKKVKKKHL-----DREQLQAKHQVLRK--HQHEGTLT 108

RESULT 6
A25876
vitellogenin III precursor - fruit fly (Drosophila melanogaster)
N:Alternate names: yolk polypeptide yPIII; yolk protein 3
C:Species: Drosophila melanogaster
C:Date: 04-Mar-1988 #sequence_revision 12-May-1994 #text_change 20-Aug-1999
C:Accession: A25876; A27388
R:Yan, Y.L.; Kunert, C.J.; Postlethwait, J.H.
Nucleic Acids Res. 15, 67-85, 1987
A:Title: Sequence homologies among the three yolk polypeptide (Yp) genes in Drosophila
A:Reference number: A25876; MUID:87146365
A:Accession: A25876
A:Molecule type: DNA
A:Residues: 1-420 <YAN>
A:Cross-references: GB:X04754; NID:g9844; PIDN:CAA28451.1; PID:g9845
R:Garabedian, M.J.; Shirras, A.D.; Bowles, M.; Wensink, P.C.
Gene 55, 1-8, 1987
A:Title: The nucleotide sequence of the gene coding for Drosophila melanogaster yolk protein 3
A:Reference number: A27388; MUID:87305580
A:Accession: A27388
A:Molecule type: DNA
A:Residues: 1-420 <GAR>
A:Cross-references: GB:M15898; NID:g158815; PIDN:AAA29024.1; PID:g158816
C:Genetics:
A:Gene: FlyBase:Yp3
A:Cross-references: FlyBase:FBgn0004047
C:Superfamily: Insect vitellogenin
C:Keywords: egg yolk; glycoprotein; phosphoprotein; sulfoprotein
F;1-19/Domain: signal sequence #status predicted <SIG>

Query Match 9.0%; Score 76.5; DB 2; Length 420;
Best Local Similarity 24.4%; Pred. No. 16;
Matches 39; Conservative 14; Mismatches 54; Indels 53; Gaps 7;

Qy 8 MERIVICLMVIFGLTVHKSSQGDQ-----HMIRMRLIDI-----45
Db 1 MMSLRICLATTCLLVAHAASKDASNDRLKPTKWLATATELENVPSLNDITWERLENQPLSQ 60
Qy 46 -----VDQLKNYNDLVPEFLPAPEDVETNCWSAFSCFQKAQLK-----SA 87
Db 61 CAKVIEKIVHVGQIK---HDLTPSFVPSPNPV---WIKSNGQKVECKLANNYVETAKA 114
Qy 88 NFGNNERIINVSIIKKL-KRKPPSTNAGRR-----OKHRL 120
Db 115 QPFGGEDEVTIVLTGLPTSPAQQAMRRLIOAYVOKYNL 154

RESULT 7
S60441
hypothetical protein YGR150c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein G6642
C:Species: Saccharomyces cerevisiae
C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 29-Oct-1999
C:Accession: S60441; S64459
R:Skala, J.; Nawrocki, A.; Goffeau, A.
Yeast 11, 1421-1427, 1995
A:Title: The sequence of a 27 kb segment on the right arm of chromosome VII from Sacc
A:Reference number: S60435; MUID:96158062
A:Accession: S60441
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-864 <SKA>
A:Cross-references: EMBL:X85807; NID:g1045249; PIDN:CAA59808.1; PID:g1045256
R:Van Dyck, L.; Skala, J.; de Wergifosse, P.; Purnelle, B.; Talla, E.; Nawrocki, A.;
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64428
A:Accession: S64459
A:Molecule type: DNA
A:Residues: 1-864 <VAN>
A:Cross-references: EMBL:Z72935; NID:g1323253; PIDN:CAA97164.1; PID:e243709; PID:g132
A:Experimental source: strain S288C
C:Genetics:
A:Map position: 7R

Query Match 9.0%; Score 76.5; DB 2; Length 864;
Best Local Similarity 22.5%; Pred. No. 34;
Matches 43; Conservative 22; Mismatches 59; Indels 67; Gaps 8;

Qy 2 RSSPGNMRIVIC-----LMVIFGLTVHKSSQGDQDRHMRMRLIDIVDLKKNYNDLV 57
Db 5 RCGPRNN---VLCFPQLSFLSKRLINK-----RFKYTLQTEDE-KNMGSL 49
Qy 58 PEFPLPAPEDVETNCWSAFSCFQKAQLKSANTGNERRIINVSIIKKLRKPPSTNAGRRQK 117
Db 50 KKKIITPEDVE-----FKLAQLREFSNTLKERIHNTK-----SYNSDGHQS 90
Qy 118 HRLTSPSCDSYEKKPPK-----EFLERFKSLQKMIHQ-----150
Db 91 NSIAPISEDNRVNVVTKTSVPNEEKSKNLSLIHSSFLKMDHLVPKVIKRVIRVADDDIL 150
Qy 151 --HLSRSTHGS 159
Db 151 AKNLFDPSHN 161

RESULT 8
S42022
ureidoglycolate hydrolase (EC 3.5.3.19) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YIR032c

R;Klemenz, R.; Hoffmann, S.; Werenskiold, A.K.
Proc. Natl. Acad. Sci. U.S.A. 86, 5708-5712, 1989
A:Title: Serum- and oncoprotein-mediated induction of a gene with sequence similarity to
A:Reference number: A33541; MUID:89345536
A:Accession: A33541
A:Molecule type: mRNA
A:Residues: 1-191,'V',193-328,'SKECPSHIA' <KLE>
A:CROSS-references: GB:I24843; NID:g201103; PIDN:AAA40160.1; PID:g201104
R:Tominaga, S.I.; Jenkins, N.A.; Gilbert, D.J.; Copeland, N.G.; Tetsuka, T.
Biochim. Biophys. Acta 1090, 1-8, 1991
A:Title: Molecular cloning of the murine ST2 gene. Characterization and chromosomal mapping
A:Reference number: S17657; MUID:91355215
A:Accession: S17657
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-191,'V',193-328,'SKECPSHIA' <TOM>
A:CROSS-references: EMBL:X60184; NID:g54200; PIDN:CAA42742.1; PID:g54201
R:Tominaga, S.I., 1993
FEBS Lett. 258, 301-304, 1989
A:Title: A putative protein of a growth specific cDNA from BALB/c-3T3 cells is highly similar
A:Reference number: S07054; MUID:90092495
A:Accession: S07054
A:Molecule type: mRNA
A:Residues: 1-328,'SKECPSHIA' <TO2>
A:CROSS-references: EMBL:X07519; NID:g55517; PIDN:CAA68812.1; PID:g55518
A:Note: it is uncertain whether Met-1, Met-7 or Met-19 is the initiator
C:Genetics:
A:Gene: ST2
A:Map position: 1
A:Introns: 27/1; 97/2; 155/3; 210/1; 233/1; 280/2
C:Superfamily: interleukin-1 receptor type I
C:Keywords: glycoprotein
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-337/Product: ST2 protein #status predicted <MAT>
F:60,101,107,146,176,194,225,259,278/Binding site: carbohydrate (Asn) (covalent) #status

Query Match 10.8%; Score 91.5; DB 2; Length 567;
Best Local Similarity 25.6%; Pred. No. 0.9;
Matches 33; Conservative 20; Mismatches 39; Indels 37; Gaps 7;

QY 47 DOLKNVNDLVPEFLPAPEDVFNCWEASFCFOKAQLKSANTGNNERIINVSIKLKRRK 106
||| |||| : : : : :
Db 76 DLKK-----FLPARVE-----DSGIYACV----IRSPNL-NKTGYLNVTIHK---K 113
||| | : : : : :
QY 107 PPSTN-----AGRQKHRLTPSCDSYEKKPKPELFERFKSLLOQMIOH----- 151
||| | : : : : :
Db 114 PPSCNIPDYLMTVGSKDNFTICTIDLYNWTPQWFKNKALQEPFRAHRSYLF 173
||| | : : : : :
QY 152 LSSRTGHSSE 160
: : |||
Db 174 IDNVTHDE 182

RESULT 3
B38529
nikB protein - Escherichia coli plasmid R64
C:Species: Escherichia coli
C>Date: 12-Jul-1991 #sequence_revision 12-Jul-1991 #text_change 28-Jul-2000
C:Accession: B38529
R:Furuya, N.; Nisiooka, T.; Komano, T.
J. Bacteriol. 173, 2231-2237, 1991
A:Title: Nucleotide sequence and functions of the *oriT* operon in IncII plasmid R64.
A:Reference number: A38529; MUID:91177811
A:Accession: B38529
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-899 <FUR>
A:CROSS-references: GB:D90273; NID:g217081; PIDN:BAA14318.1; PID:g217083
C:Genetics:
A:Genome: plasmid
C:Superfamily: Escherichia coli nikB protein

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2001, 09:37:38 ; Search time 14.29 seconds
(without alignments)
779.084 Million cell updates/sec

Title: US-09-522-217-2

Perfect score: 850

Sequence: 1 MRSSPGNMRIVICLMVFL.....LLQKMIHQHLSRTHGSEDS 162

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_67:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94.5	11.1	162	1 A53484	interleukin-15 pre
2	91.5	10.8	567	2 S29498	lymphocyte antigen
3	83	9.8	899	2 B38529	nikB protein - Esc
4	82	9.6	607	1 ABXL72	74K albumin precu
5	79	9.3	206	2 S49882	hypothetical prote
6	76.5	9.0	420	2 A25876	vitellogenin III p
7	76.5	9.0	864	2 S60441	hypothetical prote
8	76	8.9	195	2 S42022	ureidoglycolate hy
9	75.5	8.9	336	2 S42632	Fit-1S protein pre
10	75.5	8.9	406	2 T28957	hypothetical prote
11	75.5	8.9	1206	2 S24407	formin isoform IV
12	75	8.8	321	2 T46352	hypothetical prote
13	74.5	8.8	286	2 F82881	hypothetical prote
14	74.5	8.8	482	2 S41211	voltage-dependent
15	74.5	8.8	789	2 S44759	Cl14B9.5 protein -
16	74.5	8.8	848	2 T00372	hypothetical prote
17	74	8.7	440	2 T47906	FUSCA PROTEIN FUS6
18	74	8.7	953	2 I48078	CHO1 antigen - Chl
19	74	8.7	2197	2 B71600	variant-specific s
20	73.5	8.6	162	2 I49124	interleukin-15 - m
21	73.5	8.6	447	2 S39316	CAB3B protein - hu
22	73.5	8.6	472	2 I59087	ISG-K54 - human
23	73.5	8.6	477	2 S21049	calcium channel pr
24	73.5	8.6	484	2 A4608	voltage-dependent
25	73.5	8.6	484	2 S39315	CAB3a protein - hu
26	73.5	8.6	868	2 T31527	hypothetical prote
27	73.5	8.6	2118	2 T13612	hypothetical prote
28	73	8.6	304	2 A32108	translation initia
29	73	8.6	1291	2 T22382	hypothetical prote

ALIGNMENTS

RESULT 1

A53484

interleukin-15 precursor - green monkey

C:Species: Cercopithecus aethiops (green monkey, grivet)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A53484

R:Grabstein, K.H.; Eisenman, J.; Shanebeck, K.; Rauch, C.; Srinivasan, S.; Fung, V.; Giri, J.G.

Science 264, 965-968, 1994

A:Title: Cloning of a T cell growth factor that interacts with the beta chain of the

A:Reference number: A53484; MUID:94233380

A:Accession: A53484

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-162 <GRA>

A:Cross-references: GB:U03099; NID:9493521; PIDN:AAA18416.1; PID:9493522

A:Note: the complete translation is not shown

C:Superfamily: interleukin-15

C:Keywords: growth factor

F:49-162/Product: interleukin-15 #status predicted <MAT>

F:83-133,90-136/Disulfide bonds: #status predicted

Query Match 11.1%; Score 94.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.13;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

Oy 43 IDVDQLKNYNDLYPEF-----LPAPEDVETNCWSAFSCF-----QKAQLKSANTGNNE 93

Db 51 VNVISDLKK-IEDLIQSMHIDATLTESDVHPSCVKVTAMKCFLELQVISHESGDDIHD 109

Oy 94 RLINSYIKKKRKPPSTNAGRQKHRLTGPCSDSYVEKKPKPFLEFKSLLOKMIH 149

Db 110 TVENLI--LANNILSSNGNITES---GCKECELEEKNIKEFLQSFVHIVQMFN 160

RESULT 2

S29498

lymphocyte antigen Ly84 precursor - mouse

N:Alternate names: 38.5K T1 glycoprotein; ST2L protein

C:Species: Mus musculus (house mouse)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Dec-2000

C:Accession: S29498; A33541; S17657; S07054

R:Yanagisawa, K.; Takagi, T.; Tsukamoto, T.; Tetsuka, T.; Tominaga, S.

FEBS Lett. 318, 83-87, 1993

A:Title: Presence of a novel primary response gene ST2L, encoding a product highly si

A:Reference number: S29498; MUID:93170492

A:Accession: S29498

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-567 <YAN>

A:Cross-references: EMBL:D13695; NID:g286100; PIDN:BAA02854.1; PID:g286101

✓

TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-587-0430
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-535-733-2

Query Match 11.7%; Score 99.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.0011;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

Oy 43 IDIVDLKKNYNDLVPEF-----LPAPEDVETNCWSAFSCF-----OKAQLKSANTGNNE 93
Db 51 VNVISDLKK-IEDLIQSMHIDATLVTSVDHPSCKVTAMKCFLELQVISLESGDASIHD 109
Oy 94 RIINVSIIKKLRKPPSTNAGRQKHRLTGPCSDSYEKKPPKPEFLERFKSLQKMIH 149
Db 110 TVENLII--LANNSSNGNVTES--GCKECELEEKNIKEFLQSFVHIQVOMFIN 160

RESULT 15
US-08-726-817-2
Sequence 2, Application US/08726817
Patent No. 5707616
GENERAL INFORMATION:
APPLICANT: Grabstein, Kenneth
APPLICANT: Anderson, Dirk
APPLICANT: Eisenman, June
APPLICANT: Fung, Victor
APPLICANT: Rauch, Charles
TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,817
FILING DATE: 04-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/393,305
FILING DATE: 22-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 480052.409C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-726-817-2

Query Match 11.7%; Score 99.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.0011;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

Oy 43 IDIVDLKKNYNDLVPEF-----LPAPEDVETNCWSAFSCF-----OKAQLKSANTGNNE 93
Db 51 VNVISDLKK-IEDLIQSMHIDATLVTSVDHPSCKVTAMKCFLELQVISLESGDASIHD 109
Oy 94 RIINVSIIKKLRKPPSTNAGRQKHRLTGPCSDSYEKKPPKPEFLERFKSLQKMIH 149
Db 110 TVENLII--LANNSSNGNVTES--GCKECELEEKNIKEFLQSFVHIQVOMFIN 160

Search completed: May 24, 2001, 09:37:55
Job time: 17 sec

Db 62 TVENLIILANNLSSNGNXITESG-----CKECELEEKNIKEFLQSFVHIVQMFN 112

RESULT 12

US-08-031-399-5

; Sequence 5, Application US/08031399

; Patent No. 5552303

; GENERAL INFORMATION:

; APPLICANT: Grabstein, Kenneth

; APPLICANT: Anderson, Dirk

; APPLICANT: Eisenman, June

; APPLICANT: Fung, Victor

; APPLICANT: Rauch, Charles

; TITLE OF INVENTION: Epithelium-derived T-cell Factor

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Immunex Corporation

; STREET: 51 University Street

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98101

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/031,399

; FILING DATE: 19930308

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Launer, Charlene

; REGISTRATION NUMBER: 33,035

; REFERENCE/DOCKET NUMBER: 2811

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-587-0430

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 162 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-031-399-5

Query Match 11.7%; Score 99.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.0011;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 43 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWNSAFSCF-----OKAQKLSANTGNNE 93

Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVLSLESGDASIH 109

QY 94 RIINVSIIKKLRKPPSTNAGRRQKRLTFCPSDCSYEKKPPKEFLERKSLLOKMIH 149

Db 110 TVENLII--LANNLSSNGNVTES---GCKECELEEKNIKEFLQSFVHIVQMFN 160

RESULT 13

US-08-393-305-2

; Sequence 2, Application US/08393305

; Patent No. 5574138

; GENERAL INFORMATION:

; APPLICANT: Grabstein, Kenneth

; APPLICANT: Anderson, Dirk

; APPLICANT: Eisenman, June

; APPLICANT: Fung, Victor

; APPLICANT: Rauch, Charles

; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/393,305

; FILING DATE: 22-FEB-1995

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: McMasters, David D.

; REGISTRATION NUMBER: 33,963

; REFERENCE/DOCKET NUMBER: 480052.409C2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-622-4900

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 162 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-393-305-2

Query Match 11.7%; Score 99.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.0011;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 43 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWNSAFSCF-----OKAQKLSANTGNNE 93

Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVLSLESGDASIH 109

QY 94 RIINVSIIKKLRKPPSTNAGRRQKRLTFCPSDCSYEKKPPKEFLERKSLLOKMIH 149

Db 110 TVENLII--LANNLSSNGNVTES---GCKECELEEKNIKEFLQSFVHIVQMFN 160

RESULT 14

US-08-535-733-2

; Sequence 2, Application US/08535733

; Patent No. 5660824

; GENERAL INFORMATION:

; APPLICANT: Quinn, LeBris

; APPLICANT: Grabstein, Kenneth

; APPLICANT: Troutt, Anthony B.

; TITLE OF INVENTION: Muscle-Trophic Factor

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Immunex Corporation

; STREET: 51 University Street

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98101

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: Apple Macintosh

; OPERATING SYSTEM: System 7, Word 5.1a

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/535,733

; FILING DATE:

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Malaska, Stephen L.

; REGISTRATION NUMBER: 32,655

; REFERENCE/DOCKET NUMBER: 2833


```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/189,193
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,305
; FILING DATE: 22-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-189-193-3

Query Match 11.7%; Score 99.5; DB 4; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00069;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 43 IDVDQLKNVNDLVPEF-----LPAPEDVTNCEWSAFSCF-----QKAQLKSANTGNNE 93
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTAMKCFLLLEQLVISLESGDASIH 61
QY 94 RIINVSIIKKLRPPSTNAGRQKRLTCTPCSDSYEKKPKPEFLERFKSLLOKMIH 149
Db 62 TVENLII--LANNSSNGNVTES---GCKECEELEEKNIKEFLQSFVHIVQMF 112

RESULT 10
PCT-US94-03793-6
; Sequence 6, Application PC/TUS9403793
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Interleukin-15
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03793
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; PCT-US94-03793-6
```

```
;
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; PCT-US94-03793-6

Query Match 11.7%; Score 99.5; DB 5; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00069;
Matches 30; Conservative 23; Mismatches 46; Indels 15; Gaps 5;

QY 43 IDVDQLKNVNDLVPEF-----LPAPEDVTNCEWSAFSCF-----QKAQLKSANTGNNE 93
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTAMKCFLLLEQLVISLESGDASIH 61
QY 94 RIINVSIIKKLRPPSTNAGRQKRLTCTPCSDSYEKKPKPEFLERFKSLLOKMIH 149
Db 62 TVENLII--LANNSSNGNVTES---GCKECEELEEKNIKEFLQSFVHIVQMF 112

RESULT 11
PCT-US94-03793-12
; Sequence 12, Application PC/TUS9403793
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Interleukin-15
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03793
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; PCT-US94-03793-12

Query Match 11.7%; Score 99.5; DB 5; Length 114;
Best Local Similarity 25.2%; Pred. No. 0.00069;
Matches 30; Conservative 23; Mismatches 45; Indels 21; Gaps 5;

QY 43 IDVDQLKNVNDLVPEF-----LPAPEDVTNCEWSAFSCF-----QKAQLKSANTGNN--- 92
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTAMKCFLLLEQLVISXESGDXXIH 61
QY 93 --ERINVSIIKKLRPPSTNAGRQKRLTCTPCSDSYEKKPKPEFLERFKSLLOKMIH 149
```



```
; FILING DATE: 04-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,305
; FILING DATE: 22-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-726-817-3

Query Match 11.7%; Score 99.5; DB 1; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00069;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 43 IDIVDLKKNYNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKAQLKSANTGNNE 93
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTAMKCFLELQVLQVISLESQDASIH 61
QY 94 RIINVSIIKKLRPPSTNAGRQKHRLTCPCSDSYEKPKPKPEFLERFKSLLOKMIH 149
Db 62 TVENLII--LANNSSNGNVTES--GCKECELEEKNIKEFLQSFVHIVQMF 112

RESULT 5
US-08-504-042-6
; Sequence 6, Application US/08504042
; Patent No. 5747024
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Epithelium-derived T-cell Factor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,042
; FILING DATE: 19-JUL-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/031,399
; FILING DATE: 08-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-504-042-12

Query Match 11.7%; Score 99.5; DB 1; Length 114;
Best Local Similarity 25.2%; Pred. No. 0.00069;
Matches 30; Conservative 23; Mismatches 45; Indels 21; Gaps 5;

QY 43 IDIVDLKKNYNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKAQLKSANTGN 92
Db 62 TVENLII--LANNSSNGNVTES--GCKECELEEKNIKEFLQSFVHIVQMF 112
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-504-042-6

Query Match 11.7%; Score 99.5; DB 1; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00069;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 43 IDIVDLKKNYNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKAQLKSANTGNNE 93
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTAMKCFLELQVLQVISLESQDASIH 61
QY 94 RIINVSIIKKLRPPSTNAGRQKHRLTCPCSDSYEKPKPKPEFLERFKSLLOKMIH 149
Db 62 TVENLII--LANNSSNGNVTES--GCKECELEEKNIKEFLQSFVHIVQMF 112

RESULT 6
US-08-504-042-12
; Sequence 12, Application US/08504042
; Patent No. 5747024
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Epithelium-derived T-cell Factor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,042
; FILING DATE: 19-JUL-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/031,399
; FILING DATE: 08-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-504-042-12

Query Match 11.7%; Score 99.5; DB 1; Length 114;
Best Local Similarity 25.2%; Pred. No. 0.00069;
Matches 30; Conservative 23; Mismatches 45; Indels 21; Gaps 5;

QY 43 IDIVDLKKNYNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKAQLKSANTGN 92
Db 62 TVENLII--LANNSSNGNVTES--GCKECELEEKNIKEFLQSFVHIVQMF 112
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2001, 09:37:38 ; Search time 11.23 seconds
(without alignments)
277.129 Million cell updates/sec

Title: US-09-522-217-2

Perfect score: 850

Sequence: 1 MRSSPGNMRIVICLWIFL.....LQKMIHQHLSRTHGSEDS 162

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_AA:*
- 1: /cgn2.6/ptodata/2/iaa/5A_COMB.pep.*
 - 2: /cgn2.6/ptodata/2/iaa/5B_COMB.pep.*
 - 3: /cgn2.6/ptodata/2/iaa/6A_COMB.pep.*
 - 4: /cgn2.6/ptodata/2/iaa/6B_COMB.pep.*
 - 5: /cgn2.6/ptodata/2/iaa/PCTUS_COMB.pep.*
 - 6: /cgn2.6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99.5	11.7	114	1	US-08-031-399-6
2	99.5	11.7	114	1	US-08-031-399-12
3	99.5	11.7	114	1	US-08-393-305-3
4	99.5	11.7	114	1	US-08-726-817-3
5	99.5	11.7	114	1	US-08-504-042-6
6	99.5	11.7	114	1	US-08-504-042-12
7	99.5	11.7	114	2	US-08-725-969-3
8	99.5	11.7	114	2	US-08-794-524-3
9	99.5	11.7	114	4	US-09-189-193-3
10	99.5	11.7	114	5	PCT-US94-03793-6
11	99.5	11.7	114	5	PCT-US94-03793-12
12	99.5	11.7	162	1	US-08-031-399-5
13	99.5	11.7	162	1	US-08-393-305-2
14	99.5	11.7	162	1	US-08-535-733-2
15	99.5	11.7	162	1	US-08-726-817-2
16	99.5	11.7	162	1	US-08-504-042-5
17	99.5	11.7	162	2	US-08-725-969-2
18	99.5	11.7	162	2	US-08-794-524-2
19	99.5	11.7	162	3	US-08-842-947-6
20	99.5	11.7	162	4	US-09-189-193-2
21	99.5	11.7	162	5	PCT-US94-03793-5
22	99.5	11.7	162	5	PCT-US96-08423-2
23	94.5	11.1	114	1	US-08-031-399-3
24	94.5	11.1	114	1	US-08-393-305-6
25	94.5	11.1	114	1	US-08-726-817-6
26	94.5	11.1	114	1	US-08-504-042-3
27	94.5	11.1	114	2	US-08-725-969-6

28	94.5	11.1	114	2	US-08-794-524-6	Sequence 6, Appli
29	94.5	11.1	114	4	US-09-189-193-6	Sequence 6, Appli
30	94.5	11.1	114	5	PCT-US94-03793-3	Sequence 3, Appli
31	94.5	11.1	122	1	US-08-300-903A-3	Sequence 3, Appli
32	94.5	11.1	162	1	US-08-031-399-2	Sequence 2, Appli
33	94.5	11.1	162	1	US-08-393-305-5	Sequence 5, Appli
34	94.5	11.1	162	1	US-08-284-393B-9	Sequence 9, Appli
35	94.5	11.1	162	1	US-08-726-817-5	Sequence 5, Appli
36	94.5	11.1	162	1	US-08-504-042-2	Sequence 5, Appli
37	94.5	11.1	162	1	US-08-725-969-5	Sequence 5, Appli
38	94.5	11.1	162	2	US-08-794-524-5	Sequence 5, Appli
39	94.5	11.1	162	3	US-08-842-947-8	Sequence 8, Appli
40	94.5	11.1	162	4	US-09-189-193-5	Sequence 5, Appli
41	94.5	11.1	162	5	PCT-US94-03793-2	Sequence 2, Appli
42	94.5	11.1	162	5	PCT-US95-08950-9	Sequence 9, Appli
43	91.5	10.8	337	1	US-08-442-043A-18	Sequence 18, Appli
44	75	8.8	334	2	US-08-933-750C-16	Sequence 16, Appli
45	75	8.8	334	4	US-09-234-7513-16	Sequence 16, Appli

ALIGNMENTS

RESULT 1
US-08-031-399-6
; Sequence 6, Application US/08031399
; Patent No. 5552303
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Epithelium-derived T-cell Factor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/031,399
; FILING DATE: 19930308
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
US-08-031-399-6

Query Match 11.7%; Score 99.5; DB 1; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00069;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;
QY 43 IDVDLQKNVNDLVPEF-----LPAPEVETNCWSAFSCF----OKAQLKSANTGNNE 93
::: :::: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||:: ||::

XX The simian ETF (epithelium derived T cell factor) was isolated from
 CC African green monkey CV1/EBNA cell conditioned medium. The N-
 CC terminal sequence of the purified sETF was determined and then PCR
 CC primers were designed based on the sequence information. A 92 bp
 CC fragment was amplified from CV1/EBNA DNA and was used as a probe to
 CC screen a CV1/EBNA cDNA library for the full-length sETF coding
 CC sequence. Mature sETF induces proliferation and/or differentiation
 CC of precursor or mature T cells and is useful for promoting long-term
 CC in vitro culture of T-lymphocytes and T-cell lines. It is used for
 CC treating gastrointestinal diseases including peptic ulcer, colitis and
 CC malignancy and for treating HIV infection.
 XX
 SQ Sequence 114 AA;

Query Match 11.7%; Score 99.5; DB 17; Length 114;
 Best Local Similarity 25.94; Pred. No. 0.0016;
 Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;
 QY 43 IDIVDQKKNVNDLVPEF-----LPAPEDVETNCEWSAFSCF---OKAOLKSANTGNNE 93
 Db 3 vnvvisdikk-iedliqsmhidatlytesdvhpckvtamkcflllelqvvisiesgdasihd 61
 QY 94 RIINVSIIKKLRKPPSTNAGRRQKHRLTCPCSDSYEKKKPPKEFLERFKSLQKMIH 149
 Db 62 tvenlii--lannslssngnvtess---gckeeceeleeknikelfqsfvhiqgmfin 112

RESULT 12
 ID W07253
 AC W07253;
 DT 05-FEB-1997 (first entry)
 DE Generic mammalian epithelium-derived T cell factor.
 KW Mammalian; epithelium-derived T-cell factor; simian; human; culture;
 KW epithelial cell; proliferation; differentiation; T-lymphocyte.
 OS Synthetic.

Key Location/Qualifiers
 FT Misc-difference 52 /label= Leu, His
 FT Misc-difference 57
 FT Misc-difference 58 /label= Ala, Thr
 FT Misc-difference 58 /label= Ser, Asp
 FT Misc-difference 73 /label= Ser, Ile
 FT Misc-difference 80 /label= Val, Ile
 US5552303-A.
 03-SEP-1996.
 08-MAR-1993; 93US-0031399.
 08-MAR-1993; 93US-0031399.
 (IMMV) IMMUNEX CORP.
 Anderson D, Eisenman J, Fung V, Grabstein K, Rauch C;
 WPI; 1996-412063/41.
 New isolated simian and human epithelium-derived T-cell factors -
 PT which stimulate the proliferation and/or differentiation of
 PT T-lymphocytes and T-cell lines

XX Claim 1; Column 29; 22pp; English.
 PS
 CC This is the amino acid sequence of a generic mammalian epithelium-derived
 CC T cell factor (ETF). The pref. ETF are the simian (W07254) or human
 CC (W07255) proteins. This sequence represents the generic sequence of the
 CC mature protein. Both the human and simian proteins contain a 48 amino
 CC acid leader sequence. ETF is a protein of 15-17 kD which is expressed by
 CC epithelial cells and stimulates proliferation and/or differentiation of
 CC precursor and/or mature T cells. The protein is therefore useful for
 CC promoting long term in vivo culture of T-lymphocytes and T-cell lines.
 XX
 SQ Sequence 114 AA;

Query Match 11.7%; Score 99.5; DB 17; Length 114;
 Best Local Similarity 25.2%; Pred. No. 0.0016;
 Matches 30; Conservative 23; Mismatches 45; Indels 21; Gaps 5;
 QY 43 IDIVDQKKNVNDLVPEF-----LPAPEDVETNCEWSAFSCF---OKAOLKSANTGN--- 92
 Db 3 vnvvisdikk-iedliqsmhidatlytesdvhpckvtamkcflllelqvvisiesgdxxihd 61
 QY 93 --ERIINVSIIKKLRKPPSTNAGRRQKHRLTCPCSDSYEKKKPPKEFLERFKSLQKMIH 149
 Db 62 tvenliilannslssngnxtesg-----ckeeceeleeknikelfqsfvhiqgmfin 112

RESULT 13
 ID W39186
 AC W39186;
 DT 08-MAY-1998 (first entry)
 DE Simian epithelium derived T-cell factor mature protein.
 KW Epithelium derived T-cell factor; ETF; simian; gastrointestinal disease;
 KW B-cell stimulation; epithelium; villi; CD4+; CD8+; T-cell stimulation;
 KW treatment; prevention.

Key Location/Qualifiers
 FT Protein 1..114
 FT /label= ETF
 US5707616-A.
 13-JAN-1998.
 04-OCT-1996; 96US-0726817.
 22-FEB-1995; 95US-0393305.
 08-MAR-1993; 93US-0031399.
 22-APR-1994; 94US-0233606.
 (IMMV) IMMUNEX CORP.
 Anderson DM, Eisenman JR, Fung V, Grabstein KH;
 Rauch C;
 WPI; 1998-100295/09.
 N-PSDB; V02873.
 Treatment or prevention of gastrointestinal diseases - by
 PT administering epithelium-derived T-cell factor polypeptide
 PT
 PS Claim 1a; Column 37-38; 34pp; English.
 CC This sequence represents a simian epithelium-derived T-cell factor (ETF)
 CC mature protein which is used in a method for treating or preventing

XX Claim 1; Page 33; 48pp; English.
 XX A simian species of IL-15 (sIL-15) was purified and its AA
 CC sequence and cDNA sequence analysed (see R83309, R83436,
 CC T00524, T00525). Both the simian and the human ORFs encode
 CC a precursor polypeptide (R83436, R83438). The precursor
 CC polypeptides each comprise a 48-AA leader sequence and a sequence
 CC encoding mature simian or human IL-15 polypeptides. The active
 CC simian and human IL-15 polypeptides are disclosed in R83309 &
 CC R83310 respectively. The invention also comprises other mammalian
 CC IL-15, including human IL-15, that hybridise to probes defined by
 CC R83438. A plasmid contg. a recombinant clone of human IL-15
 CC cDNA was deposited with ATCC on Feb. 19 1993 under ATCC 69245.
 CC The deposit was named 141-hETF. R83435 is a mammalian mature
 CC IL-15 polypeptide. It is a generic sequence which encompasses both
 CC R83309 (simian) and R83310 (human) IL-15 mature polypeptides.
 XX SQ Sequence 114 AA;

Query Match 11.7%; Score 99.5; DB 16; Length 114;
 Best Local Similarity 25.2%; Pred. No. 0.0016;
 Matches 30; Conservative 23; Mismatches 45; Indels 21; Gaps 5;
 QY 43 IDIVDLKKNVNDLVPF-----LPAPEDVETNCWAFSCF-OQAOLKSANTGNN---- 92
 Db 3 vnvvisdlkk-iedliqsmhidatlytesdvhpckvtamkcflllelqvlsxesgdxihd 61
 QY 93 --ERIIINVSIIKKLRKPPSTNAGRRQKRLTCTPCSDSYEKPKPPKEFLERFKSLQKMIH 149
 Db 62 tvenliilannxlssngnxtesg-----ckeceeleeknikelfqsfvhiqgmfin 112

RESULT 10
 ID R66928
 AC R66928;
 DT 04-SEP-1995 (first entry)
 DE Mammalian IL-15.
 KW Interleukin-15; IL-15; sIL-15; T-cell growth factor;
 KW antitumor; virucide.
 OS Mammalia.
 FH Key Location/Qualifiers
 FT Misc-difference 52 /label= Leu, His
 FT Misc-difference 57 /label= Ala, Thr
 FT Misc-difference 58 /label= Ser, Asp
 FT Misc-difference 73 /label= Ser, Ile
 FT Misc-difference 80 /label= Val, Ile
 XX 2A9402636-A.
 XX 28-DEC-1994.
 XX 18-APR-1994; 94ZA-0002636.
 XX 18-APR-1994; 94ZA-0002636.
 XX (IMMV) IMMUNEX CORP.
 PI Anderson DM, Eisenman JR, Fung V, Grabstein KH;
 PI Rauch C;

XX WPI; 1995-082473/11.
 XX New purified interleukin-15 - which induces T cell proliferation
 PT and differentiation, used for the treatment of tumours and viral
 PT infection
 XX Claim 1; Page 33; 47pp; English.
 XX Sinian and human IL-15 cDNAs (O84583-84) can be used to obtain
 CC cDNAs encoding other mammalian homologs of IL-15. A general
 CC sequence of mammalian IL-15 is claimed.
 XX SQ Sequence 114 AA;

Query Match 11.7%; Score 99.5; DB 16; Length 114;
 Best Local Similarity 25.2%; Pred. No. 0.0016;
 Matches 30; Conservative 23; Mismatches 45; Indels 21; Gaps 5;
 QY 43 IDIVDLKKNVNDLVPF-----LPAPEDVETNCWAFSCF-OQAOLKSANTGNN---- 92
 Db 3 vnvvisdlkk-iedliqsmhidatlytesdvhpckvtamkcflllelqvlsxesgdxihd 61
 QY 93 --ERIIINVSIIKKLRKPPSTNAGRRQKRLTCTPCSDSYEKPKPPKEFLERFKSLQKMIH 149
 Db 62 tvenliilannxlssngnxtesg-----ckeceeleeknikelfqsfvhiqgmfin 112

RESULT 11
 ID W09099 standard; Protein; 114 AA.
 AC W09099;
 DT 11-MAR-1997 (first entry)
 DE Simian mature epithelium derived T cell factor.
 KW SETF; African green monkey; CV1/EBNA cell; T-cell; B-cell;
 KW lymphocyte; proliferation; differentiation; gastrointestinal;
 KW HIV infection; human immunodeficiency virus.
 OS Cercopithecus aethiops.
 FH Key Location/Qualifiers
 FT Protein 1..114
 FT /label= mature_SETF
 XX US5574138-A.
 XX 12-NOV-1996.
 XX 08-MAR-1993; 93US-0031399.
 XX 22-FEB-1995; 95US-0393305.
 XX 08-MAR-1993; 93US-0031399.
 XX 22-APR-1994; 94US-0233606.
 XX (IMMV) IMMUNEX CORP.
 XX Anderson DM, Eisenman JR, Fung V, Grabstein KH;
 PI Rauch C;
 XX WPI; 1996-517923/51.
 XX N-PSDB; T49455.
 XX New epithelium derived T cell factor - induces proliferation of T
 PT and B cells, stimulates destruction of tumour and virus-infected
 PT cells and protects against toxicity, partic. for treating intestinal
 PT disease and HIV infection
 XX Claim 1; Fig 1; 35pp; English.

DR N-PSDB; 237358.
 PT Antisense compound useful for inhibiting human interleukin-15
 PT expression useful for treating diseases associated with interleukin-15
 PT expression
 XX
 PS Example 13; Column 43-44; 31pp; English.
 XX
 CC This sequence is the human interleukin-15. The invention relates to
 CC antisense compounds that are targeted to a 5' or 3' untranslated region
 CC (5'UTR or 3'UTR) of a nucleic acid molecule encoding human interleukin-15
 CC (IL-15), and inhibit the expression of human IL-15. The antisense
 CC inhibitors are useful for inhibiting expression of IL-15 in human
 CC cells or tissues in vitro, for treating humans or other animals suspected
 CC of having or being prone to a disease associated with IL-15 expression,
 CC e.g. infections, inflammation or tumours. The inhibitors can also be used
 CC for research or diagnostic purposes. Using antisense compounds
 CC specifically and effectively inhibits IL-15 function.
 XX
 SQ Sequence 135 AA;
 Query Match 12.2%; Score 103.5; DB 21; Length 135;
 Best Local Similarity 26.7%; Pred. No. 0.00069;
 Matches 31; Conservative 22; Mismatches 48; Indels 15; Gaps 5;
 OY 43 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF---OKAOLKSANTGNNE 93
 Db 24 vnvisdllk-iedllqsmhidatlytesdvhpckvktamkcflllelqlvislesgdsahd 82
 OY 94 RIINVSIIKKLRKPPSTNAGRRQKRLHRLTCPCSDSYEKPPKEFLERFKSLQKMIH 149
 Db 83 tvenlii--lannslsngnvtes---gckeeceeleknikeflqsfvhiqvmfin 133
 RESULT 8
 R83310
 ID R83310 standard; Protein; 114 AA.
 XX
 AC R83310;
 DT 02-FEB-1996 (first entry)
 XX
 XX Human interleukin-15 mature polypeptide.
 DE Interleukin-15; epithelium-derived T-cell factor; T lymphocyte.
 KW Homo sapiens.
 OS
 XX
 XX WO9527722-A.
 PN 19-OCT-1995.
 PD
 XX 06-APR-1994; 94WO-US03793.
 PF
 XX 06-APR-1994; 94WO-US03793.
 PR
 XX (IMMV) IMMUNEX CORP.
 PA
 XX Anderson DM, Eisenman JR, Fung V, Grabstein KH;
 PI Rauch C;
 XX
 XX WPI; 1995-373556/48.
 DR N-PSDB; T00527.
 XX
 XX Isolated DNA encoding polypeptide with mammalian IL-15 activity - which
 PT stimulates proliferation and differentiation of T cells, used for
 PT treating carcinoma(s), melanomas, etc. and viral infections
 XX
 PS Claim 23; Page 30; 48pp; English.
 XX
 XX A simian species of IL-15 (sIL-15) was purified and its AA
 CC sequence and cDNA sequence analysed (see R83309, R83436,

CC T00524, T00525). Both the simian and the human ORFs encode
 CC a precursor polypeptide (R83436, R83438). The precursor
 CC polypeptides each comprise a 48-AA leader sequence and a sequence
 CC encoding mature simian or human IL-15 polypeptides. The active
 CC simian and human IL-15 polypeptides are disclosed in R83309 &
 CC R83310 respectively. The invention also comprises other mammalian
 CC IL-15, including human IL-15, that hybridise to probes defined by
 CC R83438. A plasmid contg. a recombinant clone of human IL-15
 CC cDNA was deposited with ATCC on Feb. 19 1993 under ATCC 69245.
 CC The deposit was named 141-hETF. R83435 is a mammalian mature
 CC IL-15 polypeptide. It is a generic sequence which encompasses both
 CC R83309 (simian) and R83310 (human) IL-15 mature polypeptides.
 XX
 SQ Sequence 114 AA;
 Query Match 11.7%; Score 99.5; DB 16; Length 114;
 Best Local Similarity 25.9%; Pred. No. 0.0016;
 Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;
 OY 43 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF---OKAOLKSANTGNNE 93
 Db 3 vnvisdllk-iedllqsmhidatlytesdvhpckvktamkcflllelqlvislesgdsahd 61
 OY 94 RIINVSIIKKLRKPPSTNAGRRQKRLHRLTCPCSDSYEKPPKEFLERFKSLQKMIH 149
 Db 62 tvenlii--lannslsngnvtes---gckeeceeleknikeflqsfvhiqvmfin 112
 RESULT 9
 R83435
 ID R83435 standard; Protein; 114 AA.
 XX
 AC R83435;
 XX
 DT 02-FEB-1996 (first entry)
 XX
 DE Mammalian interleukin-15 mature polypeptide.
 XX Interleukin-15; epithelium-derived T-cell factor; T lymphocyte.
 KW Mammalian.
 OS
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 52 /label= L,H
 FT Misc-difference 57 /label= A,T
 FT Misc-difference 58 /label= S,D
 FT Misc-difference 73 /label= S,I
 FT Misc-difference 80 /label= V,I
 XX
 PN WO9527722-A.
 XX
 XX 19-OCT-1995.
 PD
 XX 06-APR-1994; 94WO-US03793.
 PF
 XX 06-APR-1994; 94WO-US03793.
 PR
 XX (IMMV) IMMUNEX CORP.
 PA
 XX Anderson DM, Eisenman JR, Fung V, Grabstein KH;
 PI Rauch C;
 XX
 XX WPI; 1995-373556/48.
 DR
 XX Isolated DNA encoding polypeptide with mammalian IL-15 activity - which
 PT stimulates proliferation and differentiation of T cells, used for
 PT treating carcinoma(s), melanomas, etc. and viral infections

```
XX 09-MAR-2000; 2000WO-US06067.
PF
XX
XX 09-MAR-1999; 99US-0264908.
PR
XX 11-MAR-1999; 99US-0265992.
PR
XX 01-JUL-1999; 99US-0142013.
PR
XX
XX (ZYMO ) ZYMOGENETICS INC.
PA
XX
XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX
XX WPI; 2000-565600/52.
DR
XX
XX New human cytokine, designated zalphall ligand, useful for stimulating
PT the proliferation and/or development of haematopoietic cells in vitro
PT and in vivo, and for treating tumourigenesis -
XX
XX Example 34; Page 227; 256pp; English.
PS
XX The present sequence was used to raise antibodies, and is derived from
CC a human zalphall ligand polypeptide. zalphall ligand is a cytokine.
CC The zalphall ligand is useful for stimulating the proliferation and
CC development of haematopoietic cells in vitro and in vivo. zalphall
CC ligand polynucleotides can be used as primers or probes for cloning the
CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
CC A zalphall ligand-saporin fusion toxin may be used for treating
CC leukaemias and lymphomas. Antagonists against zalphall ligand are
CC useful as research reagents for characterizing ligand-receptor
CC interaction. Antagonists are also useful for inhibiting expansion,
CC proliferation, activation and differentiation of cells involved in
CC regulating hematopoiesis. The zalphall ligand may also be used to
CC stimulate an immune response against B cell tumour, a virus, a parasite
CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
CC agonists and antibodies are also useful for the detection, diagnosis,
CC prevention, and treatment of diseases associated with a zalphall ligand
CC genetic defect.
XX
XX Sequence 40 AA;
SQ
Query Match 24.9%; Score 212; DB 21; Length 40;
Best Local Similarity 100.0%; Pred. NO. 1e-16;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 32 QDRHMIRMROLIDIVDOLKNVNDLVPEFLPAPEDVETNC 71
Db 1 qdrhmirmrqlidivdqlknyvndlvpeflpapedvetc 40
RESULT 6
B18626
ID B18626 standard; Peptide; 32 AA.
XX
XX AC B18626;
XX
XX 22-JAN-2001 (first entry)
DT
XX
XX Antigeninc peptide derived from a human zalphall ligand polypeptide.
DE
XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
XX tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
XX
XX Homo sapiens.
OS
XX WO200053761-A2.
XX
XX 14-SEP-2000.
PD
XX
XX 09-MAR-2000; 2000WO-US06067.
PF
XX 09-MAR-1999; 99US-0264908.
PR
XX 11-MAR-1999; 99US-0265992.
PR
XX
```

```
PR 01-JUL-1999; 99US-0142013.
XX
XX (ZYMO ) ZYMOGENETICS INC.
PA
XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX
XX WPI; 2000-565600/52.
DR
XX
XX New human cytokine, designated zalphall ligand, useful for stimulating
PT the proliferation and/or development of haematopoietic cells in vitro
PT and in vivo, and for treating tumourigenesis -
XX
XX Example 34; Page 227; 256pp; English.
PS
XX The present sequence was used to raise antibodies, and is derived from
CC a human zalphall ligand polypeptide. zalphall ligand is a cytokine.
CC The zalphall ligand is useful for stimulating the proliferation and
CC development of haematopoietic cells in vitro and in vivo. zalphall
CC ligand polynucleotides can be used as primers or probes for cloning the
CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
CC A zalphall ligand-saporin fusion toxin may be used for treating
CC leukaemias and lymphomas. Antagonists against zalphall ligand are
CC useful as research reagents for characterizing ligand-receptor
CC interaction. Antagonists are also useful for inhibiting expansion,
CC proliferation, activation and differentiation of cells involved in
CC regulating hematopoiesis. The zalphall ligand may also be used to
CC stimulate an immune response against B cell tumour, a virus, a parasite
CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
CC agonists and antibodies are also useful for the detection, diagnosis,
CC prevention, and treatment of diseases associated with a zalphall ligand
CC genetic defect.
XX
XX Sequence 32 AA;
SQ
Query Match 20.7%; Score 176; DB 21; Length 32;
Best Local Similarity 100.0%; Pred. NO. 8.1e-13;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 122 CPSCDSYEKKPKFLEPKSLLOKMIHOHLS 153
Db 1 cpscdsyekkpkflekfrksllqkmihghls 32
RESULT 7
Y54825
ID Y54825 standard; Protein; 135 AA.
XX
XX AC Y54825;
XX
XX 04-FEB-2000 (first entry)
DT
XX
XX Human Interleukin-15 protein sequence.
DE
XX
XX Interleukin-15; IL-15; antisense inhibitor; untranslated region; 5'UTR;
XX 3'UTR; human; infection; inflammation; tumour; therapy; diagnosis.
XX
XX Homo sapiens.
OS
XX US5985663-A.
XX
XX 16-NOV-1999.
PD
XX
XX 25-NOV-1998; 98US-0200141.
PF
XX 25-NOV-1998; 98US-0200141.
PR
XX (ISIS-) ISIS PHARM INC.
PA
XX Bennett CF, Cowser LM;
XX
XX WPI; 2000-022283/02.
DR
```

```

XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX WPI: 2000-565600/52.
XX N-PSDB; A75580.
XX New human cytokine, designated zalphall ligand, useful for stimulating
PT the proliferation and/or development of haematopoietic cells in vitro
PT and in vivo, and for treating tumourigenesis -
XX Disclosure; Page 222-223; 256pp; English.
XX The present sequence represents a mouse zalphall ligand polypeptide,
CC which is a cytokine. The zalphall ligand is useful for stimulating the
CC proliferation and development of haematopoietic cells in vitro and in
CC vivo. Zalphall ligand polynucleotides can be used as primers or probes
CC for cloning the zalphall gene. The zalphall ligand is useful for
CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
CC used for treating leukaemias and lymphomas. Antagonists against zalphall
CC ligand are useful as research reagents for characterizing ligand-receptor
CC interaction. Antagonists are also useful for inhibiting expansion,
CC proliferation, activation and differentiation of cells involved in
CC regulating hematopoiesis. The zalphall ligand may also be used to
CC stimulate an immune response against B cell tumour, a virus, a parasite
CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
CC agonists and antibodies are also useful for the detection, diagnosis,
CC prevention, and treatment of diseases associated with a zalphall ligand
CC genetic defect.
XX Sequence 146 AA;
SQ

```

```

Query Match 57.2%; Score 486; DB 21; Length 146;
Best Local Similarity 63.0%; Pred. No. 1.3e-46;
Matches 92; Conservative 23; Mismatches 31; Indels 0; Gaps 0;

```

```

OY 8 MERIVICLVIFLGLVHKSSQSDRHMRMROLIDVDQKNVNDLVPEFLPAPEDV 67
DB 1 mertivciviflgtvahnksspgpdrllirhlidiveqklyendlpellisapqdv 60
OY 68 ETNCEWSAFSCFOKAOLKSANTGNRIINVSIRKKLRKPPSTNAGRQKRLHLCPCSDS 127
DB 61 kghcehaafacfkaklkpsnpgnktfiidlvaglrriparrggkqkhiakpcscds 120
OY 128 YEKKPKFEFLERKSLLOKMIHQHLS 153
DB 121 yekrtpkfeflerlkwlqkmiqhls 146

```

```

RESULT 4
BI8628
ID BI8628 standard; Protein; 510 AA.
XX
XX AC BI8628;
XX
XX DT 22-JAN-2001 (first entry)
XX
XX DE Amino acid sequence of MBP-mouse zalphall ligand fusion in pTAP134.
XX
XX KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
XX tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
XX
XX OS Synthetic.
XX OS Mus musculus.
XX PN WO200053761-A2.
XX
XX PD 14-SEP-2000.
XX
XX PF 09-MAR-2000; 2000WO-US06067.
XX
XX PR 09-MAR-1999; 99US-0264908.

```

```

PR 11-MAR-1999; 99US-0265992.
PR 01-JUL-1999; 99US-0142013.
XX
XX PA (ZYMO ) ZYMOGENETICS INC.
XX
XX PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX WPI: 2000-565600/52.
XX N-PSDB; A75602.
XX New human cytokine, designated zalphall ligand, useful for stimulating
PT the proliferation and/or development of haematopoietic cells in vitro
PT and in vivo, and for treating tumourigenesis -
XX Example 31; Page 239-240; 256pp; English.
XX The present sequence represents a MBP-mouse zalphall ligand fusion in
CC the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
CC is useful for stimulating the proliferation and development of
CC haematopoietic cells in vitro and in vivo. Zalphall ligand
CC polynucleotides can be used as primers or probes for cloning the
CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
CC A zalphall ligand-saporin fusion toxin may be used for treating
CC leukaemias and lymphomas. Antagonists against zalphall ligand are
CC useful as research reagents for characterizing ligand-receptor
CC interaction. Antagonists are also useful for inhibiting expansion,
CC proliferation, activation and differentiation of cells involved in
CC regulating hematopoiesis. The zalphall ligand may also be used to
CC stimulate an immune response against B cell tumour, a virus, a parasite
CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
CC agonists and antibodies are also useful for the detection, diagnosis,
CC prevention, and treatment of diseases associated with a zalphall ligand
CC genetic defect.
XX Sequence 510 AA;
SQ

```

```

Query Match 46.4%; Score 394; DB 21; Length 510;
Best Local Similarity 62.0%; Pred. No. 1.3e-35;
Matches 75; Conservative 19; Mismatches 27; Indels 0; Gaps 0;

```

```

OY 33 DRHMRMROLIDVDQKNVNDLVPEFLPAPEDVETNCEWSAFSCFOKAOLKSANTGN 92
DB 390 drllirhlidiveqklyendlpellisapqdvkghcehaafacfkaklkpsnpgm 449
OY 93 ERIINVSIRKKLRKPPSTNAGRQKRLHLCPCSDSYEKKPKFEFLERKSLLOKMIHQH 152
DB 450 ktfiidlvaglrriparrggkqkhiakpcscdsyekrtpkfeflerlkwlqkmiqh 509
OY 153 S 153
DB 510 s 510

```

```

RESULT 5
BI8625
ID BI8625 standard; Peptide; 40 AA.
XX
XX AC BI8625;
XX
XX DT 22-JAN-2001 (first entry)
XX
XX DE Antigenic peptide derived from a human zalphall ligand polypeptide.
XX KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
XX tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
XX
XX OS Homo sapiens.
XX PN WO200053761-A2.
XX
XX PD 14-SEP-2000.

```

PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX
 XX Disclosure; Page 205-206; 256pp; English.
 XX
 CC The present sequence represents a human zalphall ligand polypeptide,
 CC which is a cytokine. The zalphall ligand is useful for stimulating the
 CC proliferation and development of haematopoietic cells in vitro and in
 CC vivo. Zalphall ligand polynucleotides can be used as primers or probes
 CC for cloning the zalphall gene. The zalphall ligand is useful for
 CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
 CC used for treating leukaemias and lymphomas. Antagonists against zalphall
 CC ligand are useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 SQ Sequence 162 AA;

Query Match 100.0%; Score 850; DB 21; Length 162;
 Best Local Similarity 100.0%; Pred. No. 2.9e-87;
 Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRSSPGNERIVICLWVIFGLTVHKSSSQDRHMRMROLIDIVDLKKNYNDLVPEF 60
 DB 1 mrsspgneriviclmvifgltlvhksssgqdrhmrmrqldivdqlknyndlvpef 60
 QY 61 LPAPEDVETNCWSAFSCFOKAQLKSANTGNNERIINVSIKKLRKPPSTNAGRQKRL 120
 DB 61 lpapedvetncwsafscfkaqlksantgnneriinvsikkkrppstnagrqrkrl 120
 QY 121 TCPSCDSYEKKPKPEFLERFKSLQKMIHQHLSRTHGSEDS 162
 DB 121 tcpscdsyekppkeflerfklskqmhqhlslsrthgseds 162

RESULT 2
 B18627
 ID B18627 standard; Protein; 519 AA.
 XX AC B18627;
 XX
 XX 22-JAN-2001 (first entry)
 XX
 XX Amino acid sequence of MBP-human zalphall ligand fusion in pTAP126.
 XX
 XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX
 XX Synthetic.
 OS Homo sapiens.
 OS
 XX WO200053761-A2.
 PN
 XX 14-SEP-2000.
 XX
 XX 09-MAR-2000; 2000WO-US06067.
 XX
 XX 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillion SR, Hammond AK;
 XX

DR WPI; 2000-565600/52.
 DR N-PSDB; A75599.

XX New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX
 XX Example 31; Page 233-235; 256pp; English.

XX The present sequence represents a MBP-human zalphall ligand fusion in
 CC the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
 CC is useful for stimulating the proliferation and development of
 CC haematopoietic cells in vitro and in vivo. Zalphall ligand
 CC polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.

XX Sequence 519 AA;

Query Match 81.8%; Score 695.5; DB 21; Length 519;
 Best Local Similarity 95.7%; Pred. No. 2.5e-69;
 Matches 134; Conservative 1; Mismatches 2; Indels 3; Gaps 1;

QY 23 LVHKSSSQDRHMRMROLIDIVDLKKNYNDLVPEFLPAPEDVETNCWSAFSCFOKA 82
 DB 23 lvhksdqdrhmrmrolidivdlkknynvdlvpeflpapedvetncwsafscfka 82
 QY 383 LVPRGS---QDRHMRMRLDIVDQLKNYNDLVPEFLPAPEDVETNCWSAFSCFOKA 439
 DB 383 lvprgs---qdrhmrmrqldivdqlknyndlvpeflpapedvetncwsafscfka 439
 QY 83 QLKSAANTGNNERIINVSIKKLRKPPSTNAGRQKRLTCPSCDSYEKKPKPEFLERFKS 142
 DB 83 qlksantgnneriinvsikkkrppstnagrqrkrltcpscdsyekppkeflerfks 142
 QY 143 LIQKMIHQHLSRTHGSEDS 162
 DB 500 liqkmiqhlsrthgseds 519

RESULT 3
 B18624
 ID B18624 standard; Protein; 146 AA.
 XX AC B18624;
 XX
 XX 22-JAN-2001 (first entry)
 XX
 XX A mouse zalphall ligand polypeptide.
 XX
 XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX
 XX Mus musculus.
 OS
 XX WO200053761-A2.
 PN
 XX 14-SEP-2000.
 XX
 XX 09-MAR-2000; 2000WO-US06067.
 XX
 XX 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2001, 09:37:38 ; Search time 19.88 seconds
(without alignments)
465.817 Million cell updates/sec

Title: US-09-522-217-2

Perfect score: 850

Sequence: 1 MRSPGNMERIVICLMVFL.....LLQKMIHOHSRTHGSEDS 162

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0401.*

```

1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	850	100.0	162	21	A human zalphall 1
2	695.5	81.8	519	21	Amino acid sequenc
3	486	57.2	146	21	A mouse zalphall 1
4	394	46.4	510	21	Amino acid sequenc
5	212	24.9	40	21	Antigenic peptide
6	176	20.7	32	21	Antigenic peptide
7	103.5	12.2	135	21	Human interleukin-
8	99.5	11.7	114	16	Mammalian interleu
9	99.5	11.7	114	16	Mammalian IL-15.
10	99.5	11.7	114	16	Simian mature epit
11	99.5	11.7	114	17	

12	99.5	11.7	114	17	Generic mammalian
13	99.5	11.7	114	19	Simian epithelium
14	99.5	11.7	114	20	Simian epithelium-
15	99.5	11.7	114	21	Mature simian epit
16	99.5	11.7	162	16	Human interleukin-
17	99.5	11.7	162	16	Human IL-15. Homo
18	99.5	11.7	162	17	Simian epithelium
19	99.5	11.7	162	17	Human epithelium-d
20	99.5	11.7	162	17	Human interleukin-
21	99.5	11.7	162	18	Wild-type interleu
22	99.5	11.7	162	18	Human interleukin-
23	99.5	11.7	162	19	Simian interleukin-
24	99.5	11.7	162	19	Simian epithelium
25	99.5	11.7	162	20	Simian epithelium-
26	99.5	11.7	162	21	Amino acid sequenc
27	99.5	11.7	162	21	Human interleukin-
28	99.5	11.7	162	21	Simian epithelium-
29	99.5	11.7	162	22	Human IL-15. Homo
30	94.5	11.1	114	16	Simian mature epit
31	94.5	11.1	114	17	Human mature epit
32	94.5	11.1	114	19	Human epithelium d
33	94.5	11.1	114	20	Human epithelium-d
34	94.5	11.1	114	21	Mature human epit
35	94.5	11.1	122	17	Recombinant flag s
36	94.5	11.1	162	16	Simian interleukin
37	94.5	11.1	162	16	Simian IL-15. Cer
38	94.5	11.1	162	17	Human epithelium d
39	94.5	11.1	162	17	Simian epithelium-
40	94.5	11.1	162	17	Simian interleukin
41	94.5	11.1	162	17	Mammalian interleu
42	94.5	11.1	162	18	Mutant interleukin
43	94.5	11.1	162	19	Human epithelium d
44	94.5	11.1	162	20	Human epithelium-d
45	94.5	11.1	162	21	Simian interleukin

ALIGNMENTS

RESULT 1

```

BI8623
ID   B18623 standard; Protein; 162 AA.
XX
AC   B18623:
XX
DT   22-JAN-2001 (first entry)
XX
DE   A human zalphall ligand polypeptide.
XX
KW   zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
KW   tumorigenesis; leukaemia; hematopoiesis; B cell tumour.
XX
OS   Homo sapiens.
XX
PN   WO200053761-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-US06067.
XX
PR   09-MAR-1999; 99US-0264908.
PR   11-MAR-1999; 99US-0265992.
PR   01-JUL-1999; 99US-0142013.
XX
(ZYMO ) ZYMOGENETICS INC.
PI   Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI   Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX
WPI: 2000-565600/52.
XX
N-PSDB; A75552.
XX
PT   New human cytokine, designated zalphall ligand, useful for stimulating

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:08:37 ; Search time 108.07 Seconds
(without alignments)
57.126 Million cell updates/sec

Title: US-09-522-217-2_COPY_41_148

Perfect score: 571

Sequence: 1 QLIDIVOLKNYVNDLVPDF.....EKKPPKEFLERFKSLQKMI 108

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	571	100.0	162	21	B18623 A human zalphall 1
2	571	100.0	519	21	B18627 Amino acid sequenc
3	339	59.4	146	21	B18624 A mouse zalphall 1
4	339	59.4	510	21	B18628 Amino acid sequenc
5	164	28.7	40	21	B18625 Antigenic peptide
6	147	25.7	32	21	B18626 Antigenic peptide
7	102.5	18.0	135	21	Y54825 Human Interleukin-
8	98.5	17.3	114	16	R83310 Human Interleukin-
9	98.5	17.3	114	16	R83435 Mammalian interleu
10	98.5	17.3	114	16	R66928 Mammalian IL-15.
11	98.5	17.3	114	17	W09099 Simian mature epit

12	98.5	17.3	114	17	W07253 Generic mammalian
13	98.5	17.3	114	19	W39186 Simian epithelium
14	98.5	17.3	114	20	W39186 Simian epithelium
15	98.5	17.3	114	21	Y52309 Mature simian epit
16	98.5	17.3	162	16	R83438 Human interleukin-
17	98.5	17.3	162	16	R66927 Human IL-15. Homo
18	98.5	17.3	162	17	W09098 Simian epithelium
19	98.5	17.3	162	17	W07255 Human epithelium-d
20	98.5	17.3	162	17	R98527 Human interleukin-
21	98.5	17.3	162	18	W37369 Wild-type interleu
22	98.5	17.3	162	18	W01658 Human interleukin-
23	98.5	17.3	162	19	W53878 Human interleukin-
24	98.5	17.3	162	19	W39185 Simian epithelium
25	98.5	17.3	162	20	Y03756 Amino acid sequenc
26	98.5	17.3	162	21	B18632 Human interleukin-
27	98.5	17.3	162	21	Y78595 Human interleukin-
28	98.5	17.3	162	21	Y52308 Simian epithelium-
29	98.5	17.3	162	22	B50870 Human IL-15. Homo
30	93.5	16.4	114	16	R83309 Simian interleukin
31	93.5	16.4	114	17	W09101 Human mature epith
32	93.5	16.4	114	19	W39188 Human epithelium d
33	93.5	16.4	114	20	Y03759 Human epithelium-d
34	93.5	16.4	114	21	Y52311 Mature human epith
35	93.5	16.4	122	17	R90842 Recombinant flag s
36	93.5	16.4	162	16	R83436 Simian interleukin
37	93.5	16.4	162	16	R66926 Simian IL-15. Cer
38	93.5	16.4	162	17	W09100 Human epithelium d
39	93.5	16.4	162	17	W07254 Simian epithelium-
40	93.5	16.4	162	17	R98526 Simian interleukin
41	93.5	16.4	162	17	R92798 Mammalian interleu
42	93.5	16.4	162	18	W37370 Mutant interleukin
43	93.5	16.4	162	19	W39187 Human epithelium d
44	93.5	16.4	162	20	Y03758 Human epithelium-d
45	93.5	16.4	162	21	Y78594 Simian interleukin

ALIGNMENTS

RESULT 1

B18623
ID B18623 standard; Protein; 162 AA.
XX B18623;
XX AC
XX DT 22-JAN-2001 (first entry)
XX DE A human zalphall ligand polypeptide.
XX KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
XX KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
XX OS Homo sapiens.
XX WO2000053761-A2.
XX PN
XX PD 14-SEP-2000.
XX PF 09-MAR-2000; 2000WO-US06067.
XX PR 09-MAR-1999; 99US-0264908.
XX PR 11-MAR-1999; 99US-0265992.
XX PR 01-JUL-1999; 99US-0142013.
XX (ZYMO) ZYMOGENETICS INC.
XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Gross JA;
XX Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX WPI; 2000-565600/52.
XX N-PSDB; A75552.
XX PT New human cytokine, designated zalphall ligand, useful for stimulating

PT the proliferation and/or development of haematopoietic cells in vitro
 XX and in vivo, and for treating tumourigenesis -
 PS Disclosure; Page 205-206; 256pp; English.
 XX
 CC The present sequence represents a human zalphall ligand polypeptide,
 CC which is a cytokine. The zalphall ligand is useful for stimulating the
 CC proliferation and development of haematopoietic cells in vitro and in
 CC vivo. Zalphall ligand polynucleotides can be used as primers or probes
 CC for cloning the zalphall gene. The zalphall ligand is useful for
 CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
 CC used for treating leukaemias and lymphomas. Antagonists against zalphall
 CC ligand are useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 SQ Sequence 162 AA;

Query Match 100.0%; Score 571; DB 21; Length 162;
 Best Local Similarity 100.0%; Pred. No. 1.9e-60;
 Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QLIDIVDQLKKNVNDLVPEFLPAPEDVETNCWSAFSCFQKALKSANTGNNRIINVSI 60
 DB 41 qlidivdqlkknvndlvpeflpapedvetncwsafscfqaqlksantgneriinvs 100
 QY 61 KKLKRPSTNAGRQKRLTQPCSDSYEKPKPKFLEKSLQKMI 108
 DB 101 kklkrppstnagrqrkrltqpcsdscyekppkpflefklsllqkmi 148

RESULT 2
 ID B18627
 XX B18627 standard; Protein; 519 AA.
 AC B18627;
 XX
 DT 22-JAN-2001 (first entry)
 XX
 DE Amino acid sequence of MBP-human zalphall ligand fusion in pTAP126.
 XX
 KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO200053761-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06067.
 XX
 PR 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 XX
 DR WPI: 2000-565600/52.
 DR N-PSDB; A75599.
 XX
 PT New human cytokine, designated zalphall ligand, useful for stimulating

PT the proliferation and/or development of haematopoietic cells in vitro
 XX and in vivo, and for treating tumourigenesis -
 PS Example 31; Page 233-235; 256pp; English.
 XX
 CC The present sequence represents a MBP-human zalphall ligand fusion in
 CC the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
 CC is useful for stimulating the proliferation and development of
 CC haematopoietic cells in vitro and in vivo. Zalphall ligand
 CC polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 SQ Sequence 519 AA;

Query Match 100.0%; Score 571; DB 21; Length 519;
 Best Local Similarity 100.0%; Pred. No. 8.3e-60;
 Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QLIDIVDQLKKNVNDLVPEFLPAPEDVETNCWSAFSCFQKALKSANTGNNRIINVSI 60
 DB 398 qlidivdqlkknvndlvpeflpapedvetncwsafscfqaqlksantgneriinvs 457
 QY 61 KKLKRPSTNAGRQKRLTQPCSDSYEKPKPKFLEKSLQKMI 108
 DB 458 kklkrppstnagrqrkrltqpcsdscyekppkpflefklsllqkmi 505

RESULT 3
 ID B18624
 XX B18624 standard; Protein; 146 AA.
 AC B18624;
 XX
 DT 22-JAN-2001 (first entry)
 XX
 DE A mouse zalphall ligand polypeptide.
 XX
 KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX
 OS Mus musculus.
 XX
 PN WO200053761-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06067.
 XX
 PR 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 XX
 DR WPI: 2000-565600/52.
 DR N-PSDB; A75580.
 XX
 PT New human cytokine, designated zalphall ligand, useful for stimulating

PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX
 XX Disclosure; Page 222-223; 256pp; English.
 XX
 CC The present sequence represents a mouse zalphall ligand polypeptide,
 CC which is a cytokine. The zalphall ligand is useful for stimulating the
 CC proliferation and development of haematopoietic cells in vitro and in
 CC vivo. Zalphall ligand polynucleotides can be used as primers or probes
 CC for cloning the zalphall gene. The zalphall ligand is useful for
 CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
 CC used for treating leukaemias and lymphomas. Antagonists against zalphall
 CC ligand are useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 XX Sequence 146 AA;
 SQ

Query Match 59.4%; Score 339; DB 21; Length 146;
 Best Local Similarity 60.7%; Pred. No. 8.8e-33;
 Matches 65; Conservative 17; Mismatches 25; Indels 0; Gaps 0;
 QY 2 LIDIVDLKKNVNDLVPPEFLPAPEDVETNCWESAFSCFQAKLKSANTGNRIINVSIK 61
 Db 35 LIDIVEQLKIYENDLPellisapqdvkgchcehaafacqkalkpsnpgnktfiidlva 94
 QY 62 KLKRPSTNAGRRQKRLTCPCSDSYEKPKPEFLERFKSLQKMI 108
 Db 95 qlrrrlparrgkqkhiakpcsdseyektrpkfeilerlkwllqkmi 141

RESULT 4
 B18628
 ID B18628 standard; Protein; 510 AA.
 XX
 AC B18628;
 XX
 DT 22-JAN-2001 (first entry)
 XX
 DE Amino acid sequence of MBP-mouse zalphall ligand fusion in pTAP134.
 XX
 KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX
 OS Synthetic.
 OS Mus musculus.
 XX
 XX WO200053761-A2.
 PN
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06067.
 XX
 PR 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 XX
 XX WPI; 2000-565600/52.
 DR N-PSDB; A75602.
 XX
 PT New human cytokine, designated zalphall ligand, useful for stimulating

PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX
 XX Example 31; Page 239-240; 256pp; English.
 XX
 CC The present sequence represents a MBP-mouse zalphall ligand fusion in
 CC the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
 CC is useful for stimulating the proliferation and development of
 CC haematopoietic cells in vitro and in vivo. Zalphall ligand
 CC polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 XX Sequence 510 AA;
 SQ

Query Match 59.4%; Score 339; DB 21; Length 510;
 Best Local Similarity 60.7%; Pred. No. 4.3e-32;
 Matches 65; Conservative 17; Mismatches 25; Indels 0; Gaps 0;
 QY 2 LIDIVDLKKNVNDLVPPEFLPAPEDVETNCWESAFSCFQAKLKSANTGNRIINVSIK 61
 Db 399 LIDIVEQLKIYENDLPellisapqdvkgchcehaafacqkalkpsnpgnktfiidlva 458
 QY 62 KLKRPSTNAGRRQKRLTCPCSDSYEKPKPEFLERFKSLQKMI 108
 Db 459 qlrrrlparrgkqkhiakpcsdseyektrpkfeilerlkwllqkmi 505

RESULT 5
 B18625
 ID B18625 standard; Peptide; 40 AA.
 XX
 AC B18625;
 XX
 DT 22-JAN-2001 (first entry)
 XX
 DE Antigenic peptide derived from a human zalphall ligand polypeptide.
 XX
 KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX
 OS Homo sapiens.
 XX
 XX WO200053761-A2.
 PN
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06067.
 XX
 PR 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 XX
 XX WPI; 2000-565600/52..
 DR
 XX New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro

PT and in vivo, and for treating tumourigenesis -
 PS Example 34; Page 227; 256pp; English.
 XX

CC The present sequence was used to raise antibodies, and is derived from
 CC a human zalphall ligand polypeptide. Zalphall ligand is a cytokine.
 CC The zalphall ligand is useful for stimulating the proliferation and
 CC development of haematopoietic cells in vitro and in vivo. Zalphall
 CC ligand polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX

SQ Sequence 40 AA;
 Query Match 28.7%; Score 164; DB 21; Length 40;
 Best Local Similarity 100.0%; Pred. No. 1.4e-12;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLIDIVDLKKNYNDLVPFLPAPEDVETNC 31
 DB 10 qlidivdlkknynyndlvpeflpapedvetnc 40
 |||||

RESULT 6
 B18626
 ID B18626 standard; Peptide; 32 AA.
 AC B18626;
 XX

DT 22-JAN-2001 (first entry)
 DE Antigenic peptide derived from a human zalphall ligand polypeptide.
 XX

KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX

OS Homo sapiens.
 XX

PN WO200053761-A2.
 PD 14-SEP-2000.
 XX

PF 09-MAR-2000; 2000WO-US06067.
 XX

PR 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX

PA (ZYMO) ZYMOGENETICS INC.
 XX

PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 XX

DR WPI; 2000-565600/52.
 XX

XX New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX

PS Example 34; Page 227; 256pp; English.
 XX

CC The present sequence was used to raise antibodies, and is derived from

CC a human zalphall ligand polypeptide. Zalphall ligand is a cytokine.
 CC The zalphall ligand is useful for stimulating the proliferation and
 CC development of haematopoietic cells in vitro and in vivo. Zalphall
 CC ligand polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX

SQ Sequence 32 AA;
 Query Match 25.7%; Score 147; DB 21; Length 32;
 Best Local Similarity 100.0%; Pred. No. 1.1e-10;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 CPSCDSYEKKPKPEFLERFKSLLOKMI 108
 DB 1 cpscdsyekppkeflerfksllqkmi 27
 |||||

RESULT 7
 Y54825
 ID Y54825 standard; Protein; 135 AA.
 XX

AC Y54825;
 XX

DT 04-FEB-2000 (first entry)
 DE Human Interleukin-15 protein sequence.
 XX

KW Interleukin-15; IL-15; antisense inhibitor; untranslated region; 5'UTR;
 KW 3'UTR; human; infection; inflammation; tumour; therapy; diagnosis.
 XX

OS Homo sapiens.
 XX

PN US5985663-A.
 XX

PD 16-NOV-1999.
 XX

PF 25-NOV-1998; 98US-0200141.
 XX

PR 25-NOV-1998; 98US-0200141.
 XX

PA (ISIS-) ISIS PHARM INC.
 XX

PI Bennett CF, Cowsert LM;
 XX

DR WPI; 2000-022283/02.
 DR N-PSDB; Z37358.
 XX

PT Antisense compound useful for inhibiting human interleukin-15
 PT expression useful for treating diseases associated with interleukin-15
 PT expression -
 XX

PS Example 13; Column 43-44; 31pp; English.
 XX

CC This sequence is the human interleukin-15. The invention relates to
 CC antisense compounds that are targeted to a 5' or 3' untranslated region
 CC (5'UTR or 3'UTR) of a nucleic acid molecule encoding human interleukin-15
 CC (IL-15), and inhibit the expression of human IL-15. The antisense
 CC inhibitors are useful for inhibiting expression of IL-15 in human
 CC cells or tissues in vitro, for treating humans or other animals suspected
 CC of having or being prone to a disease associated with IL-15 expression,
 CC e.g. infections, inflammation or tumours. The inhibitors can also be used

CC for research or diagnostic purposes. Using antisense compounds.
CC specifically and effectively inhibits IL-15 function.
XX
Sequence 135 AA: SO

Db	3	vnv:sdikk--ledllqsmhidatlytesdvhpvcsvtkmcflllelvlsesgdaashd61	
Qy	54	RIINVSIIKKLRKPPSTNAGRQKHRLTQPCSDSYVEKKPPKEFEREFKSLQKMI108	
Db	62	tvenliq--lannslsengvtes--cckeccealeeknifeflcsfshviamf111	

XX	02-FEB-1996 (first entry)
DT	
XX	
DE	Mammalian interleukin-15 mature polypeptide.
XX	
KW	Interleukin-15, epithelium-derived T-cell factor, T lymphocyte

FT	Misc-difference 57	/label= A, T
FT	Misc-difference 58	/label= S, D
FT	Misc-difference 73	/label= S, I
FT	Misc-difference 80	/label= v, r

PE 06-APR-1994; 94WO-US03793.
 XX
 PR
 PR 06-APR-1994; 94WO-US03793.
 XX
 XX
 XX (IMV) IMMUNEX CORP.
 XX
 PI Anderson DM, Eisenman JR, Fung V, Grabstein KH;
 PI Rauch C;
 DR WPI; 1995-373556/48.

PI Anderson DM, Eisenman JR, Fung V, Grabstein KH;
 PI Rauch C;
 XX
 XX WPI: 1995-373556/48.
 DR
 DR
 XX
 XX Isolated DNA encoding polypeptide with mammalian IL-15 activity - which
 PT stimulates proliferation and differentiation of T cells, used for
 PT treating carcinoma(s), melanomas, etc. and viral infections
 XX
 XX
 PS Claim 1; Page 33; 48pp; English.

Claim 1; Page 33; 48pp; English.

A simian species of IL-15 (sIL-15) was purified and its AA sequence and cDNA sequence analysed (see R83309, R83436, T00524, T00525). Both the simian and the human ORFs encode a precursor polypeptide (R83436, R83438). The precursor polypeptides each comprise a 48-AA leader sequence and a sequence encoding mature simian or human IL-15 polypeptides. The active simian and human IL-15 polypeptides are disclosed in R83309 & R83310 respectively. The invention also comprises other mammalian IL-15, including human IL-15, that hybridise to probes defined by R83438. A plasmid contg. a recombinant clone of human IL-15 cDNA was deposited with ATCC on Feb. 19 1993 under ATCC 69245. The deposit was named 141-HEF. R83435 is a mammalian mature IL-15 polypeptide. It is a generic sequence which encompasses both

CC
CC
CC
CC
CC
CC
CC
CC

a plasmid copy. A recombinant clone of human IL-15 cDNA was deposited with ATCC on Feb. 19 1993 under ATCC 69245. The deposit was named 14I-hEYF. R83435 is a mammalian mature IL-15 polypeptide. It is a generic sequence which encompasses both

CC R83309 (simian) and R83310 (human) IL-15 mature polypeptides.
XX
SQ Sequence 114 AA;

	Query Match	17.3%;	Score 98.5;	DB 16;	Length 114;	
	Best Local Similarity	25.4%;	Pred. No. 0.00035;			
	Matches	30;	Conservative	22;	Mismatches	45; Indels 21; Gaps 5;
<hr/>						
Qy	3 IDIVDQLKNYVNDLVPPE-	----	LPAEDVDETNEWAFSCF-QKAOLKSA	TGNN-----	52	
	:::	:	:	:	:	:
	:	:	:	:	:	:
Db	3 nvvisdikk-iedligsmhidatlytesdvhpscvktamckfllelqvlaxisegdgxxihd	61				
<hr/>						
Qy	53 --ERIIINVSIKKLARKPSTNAGRQRHRLTCPGSDSYEKPPFEFLRFKSLLQKMI	108.				
	: :	:	:	:	:	:
	:	:	:	:	:	:
Db	tvenliilannxlssngnxitesg-----ckeceeleeknikfeiqsfvhivcmfi	111				

RESULT	10
R66928	
ID	R66928 standard; Protein; 114 AA.
XX	
XX	
AC	R66928;
XX	
XX	
DT	04-SEP-1995 (first entry)
XX	
XX	
DE	Mammalian IL-15.
XX	
KW	Interleukin-15; IL-15; sIL-15; T-cell growth factor;
KW	antitumor; virucide.
XX	
XX	
OS	Mammalia.
OS	

Key	Location/Qualifiers
PH	52
FT	/label= Leu, His
FT	57
FT	Misc-difference
FT	/label= Ala, Thr
FT	58
FT	Misc-difference
FT	/label= Ser, Asp
FT	73
FT	Misc-difference
FT	/label= Ser, Ile
FT	80
FT	Misc-difference
FT	/label= Val, Ile

XX	2A9402636-A.
PN	
XX	
XX	28-DEC-1994.
PD	
XX	
XX	18-APR-1994; 94ZA-0002636.
PF	
XX	
XX	18-APR-1994; 94ZA-0002636.
PR	
XX	
XX	(IMMV) IMMUNEX CORP.
PA	
XX	
XX	Anderson DM, Eisenman JR, Fung V, Grabstein KH;
PI	Rauch C;
PI	

DR WPI; 1995-082473/11.

PT New purified interleukin-15 - which induces T cell proliferation
PT and differentiation, used for the treatment of tumours and viral
PT infection

PS Claim 1; Page 33; 47pp; English.

CC Simian and human IL-15 cDNAs (Q84583-84) can be used to obtain
CC cDNAs encoding other mammalian homologs of IL-15. A general
CC sequence of mammalian IL-15 is claimed.

Sequence 114 AA;

```

CC R83309 (simian) and R83310 (human) IL-15 mature polypeptides.
XX
SQ Sequence 114 AA;

Query Match 17.3%; Score 98.5; DB 16; Length 114;
Best Local Similarity 25.4%; Pred. No. 0.00035;
Matches 30; Conservative 22; Mismatches 45; Indels 21; Gaps 5;

QY 3 IDIVDLQKNYVNDLVPEF-----LPAPEDVETNCWSAFSCF-OKAOLKSANTGNN---- 52
   :::: || :||: | || :||: | || :||: | ||: | :||:
Db 3 vnvvisdikk-iedliqsmhidatlytesdvhpvcvkvmckfllelqvvisesgxixhd 61
   | :| :| | :|| | :|| | :|| | :|| | :|| | :|| |
QY 53 --ERIINVSIKKLKRKPPSTNAGRQRKHRLTQPCSDSYEKPKPKPEFLERFKSLQKMI 108
   | :| :| | :|| | :|| | :|| | :|| | :|| | :|| |
Db 62 tvenliilannxlssngnxtesg-----ckeeceeleeknikelfqsvhiqgmfi 111
   | :| :| | :|| | :|| | :|| | :|| | :|| | :|| |

RESULT 10
R66928
ID R66928 standard; Protein; 114 AA.
XX
AC R66928;
XX
XX 04-SEP-1995 (first entry)
XX Mammalian IL-15.
XX
XX Interleukin-15; IL-15; sIL-15; T-cell growth factor;
KW antitumor; virucide.
XX
XX Mammalia.
XX
XX Key Location/Qualifiers
FH MISC-difference 52 /label= Leu, His
FT MISC-difference 57 /label= Ala, Thr
FT MISC-difference 58 /label= Ser, Asp
FT MISC-difference 73 /label= Ser, Ile
FT MISC-difference 80 /label= Val, Ile
XX
XX ZA9402636-A.
PN
XX
XX 28-DEC-1994.
PD
XX
XX 18-APR-1994; 94ZA-0002636.
XX
XX 18-APR-1994; 94ZA-0002636.
PR
XX
XX (IMMV ) IMMUNEX CORP.
PA
XX
XX Anderson DM, Eisenman JR, Fung V, Grabstein KH;
PI Rauch C;
PI
XX WPI; 1995-082473/11.
XX
XX New purified interleukin-15 - which induces T cell proliferation
PT and differentiation, used for the treatment of tumours and viral
PT infection
PT
XX Claim 1; Page 33; 47pp; English.
PS
XX
XX Simian and human IL-15 cDNAs (Q84583-84) can be used to obtain
CC cDNAs encoding other mammalian homologs of IL-15. A general
CC sequence of mammalian IL-15 is claimed.
XX
SQ Sequence 114 AA;

```

Query Match	17.3%; Score 98.5; DB 16; Length 114;
Best Local Similarity	25.4%; Pred. No. 0.00035;
Matches	30; Conservative 22; Mismatches 45; Indels 21; Gaps 5;
Qy	3 IDIVDOLKNYNDLVPEF-----LPAPEDVETNCWFAFSCF-QKAOLKSANTGNN---- 52
Db	3 vnvisdlkk-iedliqsmhidatlytesdvhpsckvtamckflllelqvixesgdxxihd 61
Qy	53 --ERIINVSIIKKLRKPPSTNAGRRQHRLTCPSCDSVEKKPKPELTERFKSLLOKMI 108
Db	62 tvenlilannxlssngnxtesg-----ckeceeleeknlkeflqsfvhlvqmfi 111
RESULT	11
W09099	
ID	W09099 standard; Protein; 114 AA.
XX	
AC	W09099;
XX	
DT	11-MAR-1997 (first entry)
XX	
DE	Simian mature epithelium derived T cell factor.
XX	
KW	SETF; African green monkey; CV1/EBNA cell; T-cell; B-cell;
KW	lymphocyte; proliferation; differentiation; gastrointestinal;
KW	HIV infection; human immunodeficiency virus.
XX	
OS	Cercopithecus aethiops.
XX	
FH	Key Location/Qualifiers
FT	Protein 1..114
FT	/label= mature_SETF
XX	
PN	US5574138-A.
XX	
PD	12-NOV-1996.
XX	
PF	08-MAR-1993; 93US-0031399.
XX	
PR	22-FEB-1995; 95US-0393305.
PR	08-MAR-1993; 93US-0031399.
PR	22-APR-1994; 94US-0233606.
XX	
PA	(IMMV) IMMUNEX CORP.
XX	
PI	Anderson DM, Eisenman JR, Fung V, Grabstein KH;
PI	Rauch C;
XX	
DR	WPI; 1996-517923/51.
DR	N-PSDB; T49455.
XX	
PT	New epithelium derived T cell factor - induces proliferation of T
PT	and B cells, stimulates destruction of tumour and virus-infected
PT	cells and protects against toxicity, partic. for treating intestinal
PT	disease and HIV infection
XX	
PS	Claim 1; Fig 1; 35pp; English.
XX	
CC	The simian ETF (epithelium derived T cell factor) was isolated from
CC	African green monkey CV1/EBNA cell conditioned medium. The N-
CC	terminal sequence of the purified SETF was determined and then PCR
CC	primers were designed based on the sequence information. A 92 bp
CC	fragment was amplified from CV1/EBNA DNA and was used as a probe to
CC	screen a CV1/EBNA cDNA library for the full-length SETF coding
CC	sequence. Mature SETF induces proliferation and/or differentiation
CC	of precursor or mature T cells and is useful for promoting long-term
CC	in vitro culture of T-lymphocytes and T-cell lines. It is used for
CC	treating gastrointestinal diseases including peptic ulcer, colitis and
CC	malignancy and for treating HIV infection.
XX	
SO	Sequence 114 AA;

```

Best Local Similarity 25.4%; Pred. No. 0.00035;
Matches 30; Conservative 22; Mismatches 45; Indels 21; Gaps 5;

Qy 3 IDVIDQLKYNVDLVPEF-----LPAPEDVTNCEWSAFCF-QKALQKANTGN--- 52
Db 3 vnvslidkk-iedliqsmhidatlytesdvhpvscvvtamckfllelqvixesgxihd 61

Qy 53 --ERIINVSITKLRKPPSTNAGRQRKHRLTQPCSDSYEKKPKPEFLERFKSLQKMI 108
Db 62 tvenliilannxltssngxntesg-----ckeceeleeknikelfqsfvhiqgmfi 111

```

RESULT	13
W39186	
ID	W39186 standard; Protein; 114 AA.
XX	
XX	
AC	W39186;
XX	
DT	08-MAY-1998 (first entry)
XX	
DE	Simian epithelium derived T-cell factor mature protein.
XX	
KW	Epithelium derived T-cell factor; ERF; simian; gastrointestinal disease;
KW	B-cell stimulation; epithelium; villi; CD4+; CD8+; T-cell stimulation;
KW	treatment; prevention.

XX
OS
simian

XX	Key	Location/Qualifiers
XX	Protein	1..114
FT		/label= ETF
XX		
XX	US5707616-A.	
PN		
XX		
XX		
PD	13-JAN-1998.	
XX		
XX	04-OCT-1996;	96US-0726817.
PF		
XX		
XX	22-FEB-1995.	95US-0393305
DP		

PR 08-MAR-1993; 930S-0031399.
PR 22-APR-1994. 94NS-0233606

XX
DA / TMMV / TMMINEY CORD

XX
PI Anderson DM Eijonman TP Euna V Orobatoja v m.

PI
Rauch C;
yy

DR WPI; 1998-100295/09.
DR N-DEED: 402872

[illegible]PT administering epithelium-derived T-cell factor polypeptide
vv

This sequence represents a simian epithelium-derived T-cell factor (ETF) mature protein which is used in a method for treating or preventing

CC gastrointestinal disease. These polypeptides have particular application
CC in the treatment of gastrointestinal disorders associated with disruption

of the gastrointestinal epithelium or villi such as chemotherapy- and radiation-therapy induced enteritis (gut toxicity), mucositis, peptic ulcer disease, gastroenteritis and colitis, villus atrophic disorders, malignancy and inflammatory bowel disease. ETF polypeptides may also be useful in the treatment of human immunodeficiency virus (HIV) and HIV-associated disease due to their ability to stimulate CD4+ and CD8+ cells. Biologically active ETF may be used to treat a variety of other diseases or conditions where T-cell or B cell stimulation is desired.

SQ sequence 114 AA;

Query Match 17.3%; Score 98.5; DB 19; Length 114;
Best Local Similarity 26.1%; Pred. No. 0.00035;

Result No.	Query ⚡			DB	ID	Description
	Score	Match	Length			
1	98.5	17.3	114	1	US-08-031-399-6	Sequence 6, Appli
2	98.5	17.3	114	1	US-08-031-399-12	Sequence 12, Appli
3	98.5	17.3	114	1	US-08-393-305-3	Sequence 3, Appli
4	98.5	17.3	114	1	US-08-726-817-3	Sequence 3, Appli
5	98.5	17.3	114	1	US-08-504-042-6	Sequence 6, Appli
6	98.5	17.3	114	1	US-08-504-042-12	Sequence 12, Appli
7	98.5	17.3	114	2	US-08-725-969-3	Sequence 3, Appli
8	98.5	17.3	114	2	US-08-794-524-3	Sequence 3, Appli
9	98.5	17.3	114	4	US-09-189-193-3	Sequence 3, Appli
10	98.5	17.3	114	5	PCT-US94-03793-6	Sequence 6, Appli
11	98.5	17.3	114	5	PCT-US94-03793-12	Sequence 12, Appli
12	98.5	17.3	162	1	US-08-031-399-5	Sequence 5, Appli
13	98.5	17.3	162	1	US-08-393-305-2	Sequence 2, Appli
14	98.5	17.3	162	1	US-08-533-733-2	Sequence 2, Appli
15	98.5	17.3	162	1	US-08-726-817-2	Sequence 2, Appli
16	98.5	17.3	162	1	US-08-504-042-5	Sequence 5, Appli
17	98.5	17.3	162	2	US-08-725-969-2	Sequence 2, Appli
18	98.5	17.3	162	2	US-08-794-524-2	Sequence 2, Appli
19	98.5	17.3	162	3	US-08-842-947-6	Sequence 6, Appli
20	98.5	17.3	162	4	US-09-189-193-2	Sequence 2, Appli
21	98.5	17.3	162	5	PCT-US94-03793-5	Sequence 5, Appli
22	98.5	17.3	162	5	PCT-US96-06423-2	Sequence 2, Appli
23	93.5	16.4	114	1	US-08-031-399-3	Sequence 3, Appli
24	93.5	16.4	114	1	US-08-393-305-6	Sequence 6, Appli
25	93.5	16.4	114	1	US-08-726-817-6	Sequence 6, Appli
26	93.5	16.4	114	1	US-08-504-042-3	Sequence 3, Appli
27	93.5	16.4	114	2	US-08-725-969-6	Sequence 6, Appli

QY 3 IDIVDLKYNVDLVEF-----LPAPEDVETNCWSAFSCF-----QKAQLKSANTGNNE 53

Db 3 VNVISDLKK-IEDLIQSMHIDATLVTSVHPSCVKVTAMKCFLELQVLSLESGDASHD 61
QY 54 RIINVSIIKKLRKPPSTNAGRQKRLTQPCSDSYEKPKPEFLERFKSLQKMI 108
Db 62 TVENLII--LANNSSNGNVTES---GCKECEELEEKNIKEFLQSFVHVQVFI 111

RESULT 2

US-08-031-399-12
; Sequence 12, Application US/08031399
; Patent No. 5552303
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Epithelium-derived T-cell Factor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/031.399
; FILING DATE: 19930308
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-031-399-12

Query Match 17.3%; Score 98.5; DB 1; Length 114;
Best Local Similarity 25.4%; Pred. No. 0.00016;
Matches 30; Conservative 22; Mismatches 45; Indels 21; Gaps 5;

QY 3 IDVDQLKNYNDLVPEF-----LPAPEDVETNCWSAFSCF-OQAOLKSANTGNH 52
Db 3 VNVISDLKK-IEDLIQSMHIDATLVTSVHPSCVKVTAMKCFLELQVLSLESGDASHD 61
QY 53 --ERIINVSIIKKLRKPPSTNAGRQKRLTQPCSDSYEKPKPEFLERFKSLQKMI 108
Db 62 TVENLIIANNXLSNGNXTESG-----CKECEELEEKNIKEFLQSFVHVQVFI 111

RESULT 3

US-08-393-305-3
; Sequence 3, Application US/08393305
; Patent No. 5574138
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor

; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,305
; FILING DATE: 22-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-393-305-3

Query Match 17.3%; Score 98.5; DB 1; Length 114;
Best Local Similarity 26.1%; Pred. No. 0.00016;
Matches 30; Conservative 22; Mismatches 48; Indels 15; Gaps 5;

QY 3 IDVDQLKNYNDLVPEF-----LPAPEDVETNCWSAFSCF-OQAOLKSANTGNH 53
Db 3 VNVISDLKK-IEDLIQSMHIDATLVTSVHPSCVKVTAMKCFLELQVLSLESGDASHD 61
QY 54 RIINVSIIKKLRKPPSTNAGRQKRLTQPCSDSYEKPKPEFLERFKSLQKMI 108
Db 62 TVENLII--LANNSSNGNVTES---GCKECEELEEKNIKEFLQSFVHVQVFI 111

RESULT 4

US-08-726-817-3
; Sequence 3, Application US/08726817
; Patent No. 5707616
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,817

;; FILING DATE: 04-OCT-1996
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/393,305
;; FILING DATE: 22-FEB-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McMasters, David D.
;; REGISTRATION NUMBER: 33,963
;; REFERENCE/DOCKET NUMBER: 480052.409C2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-622-4900
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 114 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-726-817-3

Query Match 17.3%; Score 98.5; DB 1; Length 114;
Best Local Similarity 26.1%; Pred. No. 0.00016;
Matches 30; Conservative 22; Mismatches 48; Indels 15; Gaps 5;

QY 3 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-----OKAOLKSANTGNNE 53
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCKVTAMKCFLELQVLSLESGDASIH 61

QY 54 RIINVSTKKLKRKPPSTNAGRQKHRLTCPCSDSYEKKPKPEFLERKSLLOKMI 108
Db 62 TVENLII--LANNSSNGNVTES---GCKECELEEKNKEFLQSFVHVQMF 111

RESULT 5
US-08-504-042-6
; Sequence 6, Application US/08504042
; Patent No. 5747024
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Epithelium-derived T-cell Factor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,042
; FILING DATE: 19-JUL-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/031,399
; FILING DATE: 08-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids

;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: NO
US-08-504-042-6

Query Match 17.3%; Score 98.5; DB 1; Length 114;
Best Local Similarity 26.1%; Pred. No. 0.00016;
Matches 30; Conservative 22; Mismatches 48; Indels 15; Gaps 5;

QY 3 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-----OKAOLKSANTGNNE 53
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCKVTAMKCFLELQVLSLESGDASIH 61

QY 54 RIINVSTKKLKRKPPSTNAGRQKHRLTCPCSDSYEKKPKPEFLERKSLLOKMI 108
Db 62 TVENLII--LANNSSNGNVTES---GCKECELEEKNKEFLQSFVHVQMF 111

RESULT 6
US-08-504-042-12
; Sequence 12, Application US/08504042
; Patent No. 5747024
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Epithelium-derived T-cell Factor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,042
; FILING DATE: 19-JUL-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/031,399
; FILING DATE: 08-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-504-042-12

Query Match 17.3%; Score 98.5; DB 1; Length 114;
Best Local Similarity 25.4%; Pred. No. 0.00016;
Matches 30; Conservative 22; Mismatches 45; Indels 21; Gaps 5;

QY 3 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-----OKAOLKSANTGNN--- 52
Db 62 TVENLII--LANNSSNGNVTES---GCKECELEEKNKEFLQSFVHVQMF 111

Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVISESGDXIHD 61
QY 53 --ERINVISIKLKKPSTWAGRRQKRLITCPCSDSYEKPKPEFLERFKSLLOKMI 108
Db 62 TVENLIILANNXLSNGNKTESG-----CKECELEEKNIKEFLQSFVHVQMF 111

RESULT 7
US-08-725-969-3
; Sequence 3, Application US/08725969
; Patent No. 5892001
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/725.969
; FILING DATE: 04-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393.305
; FILING DATE: 22-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-725-969-3

Query Match 17.3%; Score 98.5; DB 2; Length 114;
Best Local Similarity 26.1%; Pred. No. 0.00016;
Matches 30; Conservative 22; Mismatches 48; Indels 15; Gaps 5;

QY 3 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKAQLKSANTGNNE 53
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVISESGDXIHD 61

QY 54 RIINVISIKLKKPSTWAGRRQKRLITCPCSDSYEKPKPEFLERFKSLLOKMI 108
Db 62 TVENLIILANNXLSNGNKTESG-----CKECELEEKNIKEFLQSFVHVQMF 111

RESULT 8
US-08-794-524-3
; Sequence 3, Application US/08794524
; Patent No. 5985262
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk

; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/794.524
; FILING DATE: 03-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393.305
; FILING DATE: 22-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-794-524-3

Query Match 17.3%; Score 98.5; DB 2; Length 114;
Best Local Similarity 26.1%; Pred. No. 0.00016;
Matches 30; Conservative 22; Mismatches 48; Indels 15; Gaps 5;

QY 3 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKAQLKSANTGNNE 53
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVISESGDXIHD 61

QY 54 RIINVISIKLKKPSTWAGRRQKRLITCPCSDSYEKPKPEFLERFKSLLOKMI 108
Db 62 TVENLIILANNXLSNGNKTESG-----CKECELEEKNIKEFLQSFVHVQMF 111

RESULT 9
US-09-189-193-3
; Sequence 3, Application US/09189193
; Patent No. 6184359
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/189,193
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/393,305
FILING DATE: 22-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: McMasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 480052.409C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 114 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-189-193-3

Query Match 17.3%; Score 98.5; DB 4; Length 114;
Best Local Similarity 26.1%; Pred. No. 0.00016;
Matches 30; Conservative 22; Mismatches 48; Indels 15; Gaps 5;
QY 3 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWWSAFSCF-----QKAQLKSANTGNNE 53
Db 3 VNVISDLKK-IEDLQSMHIDATLYTSDVHPSCKVTAMKCFLEQLQVISLESQDASIH 61
QY 54 RIINVSIKKLRKPPSTNAGRQKRLTQPCSDSYEKKPKPEFLERFKSLQKMI 108
Db 62 TVENLII--LANNLSNGNVTES---GCKECELEEKNIKEFLQSFVHVQMF 111

RESULT 10
PCT-US94-03793-6
Sequence 6, Application PC/TUS9403793
GENERAL INFORMATION:
APPLICANT: Grabstein, Kenneth
APPLICANT: Anderson, Dirk
APPLICANT: Eisenman, June
APPLICANT: Fung, Victor
APPLICANT: Rauch, Charles
TITLE OF INVENTION: Interleukin-15
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/03793
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Launer, Charlene
REGISTRATION NUMBER: 33,035
REFERENCE/DOCKET NUMBER: 2811
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-587-0430
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:

LENGTH: 114 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
PCT-US94-03793-6

Query Match 17.3%; Score 98.5; DB 5; Length 114;
Best Local Similarity 26.1%; Pred. No. 0.00016;
Matches 30; Conservative 22; Mismatches 48; Indels 15; Gaps 5;
QY 3 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWWSAFSCF-----QKAQLKSANTGNNE 53
Db 3 VNVISDLKK-IEDLQSMHIDATLYTSDVHPSCKVTAMKCFLEQLQVISLESQDASIH 61
QY 54 RIINVSIKKLRKPPSTNAGRQKRLTQPCSDSYEKKPKPEFLERFKSLQKMI 108
Db 62 TVENLII--LANNLSNGNVTES---GCKECELEEKNIKEFLQSFVHVQMF 111

RESULT 11
PCT-US94-03793-12
Sequence 12, Application PC/TUS9403793
GENERAL INFORMATION:
APPLICANT: Grabstein, Kenneth
APPLICANT: Anderson, Dirk
APPLICANT: Eisenman, June
APPLICANT: Fung, Victor
APPLICANT: Rauch, Charles
TITLE OF INVENTION: Interleukin-15
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/03793
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Launer, Charlene
REGISTRATION NUMBER: 33,035
REFERENCE/DOCKET NUMBER: 2811
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-587-0430
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 114 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
PCT-US94-03793-12

Query Match 17.3%; Score 98.5; DB 5; Length 114;
Best Local Similarity 25.4%; Pred. No. 0.00016;
Matches 30; Conservative 22; Mismatches 45; Indels 21; Gaps 5;
QY 3 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWWSAFSCF-----QKAQLKSANTGN-- 52
Db 3 VNVISDLKK-IEDLQSMHIDATLYTSDVHPSCKVTAMKCFLEQLQVISLESQDASIH 61
QY 53 --ERIINVSIKKLRKPPSTNAGRQKRLTQPCSDSYEKKPKPEFLERFKSLQKMI 108

Query Match	17.38;	Score 98.5;	DB 1;	Length 162;
-------------	--------	-------------	-------	-------------

Best local similarity 20.16, P-adj. NO. 0.00020,
Matches 30; Conservative 22; Mismatches 48; Indels 15; Gaps 5;

RESULT 15
US-08-726-817-2
; Sequence 2, Application US/08726817
; Patent NO. 5707616
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:

[illegible]

```

CLASSIFICATION: 433
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/393,305
  FILING DATE: 22-FEB-1995
  ATTORNEY/AGENT INFORMATION:
    NAME: McMasters, David D.
    REGISTRATION NUMBER: 33,963
    REFERENCE/DOCKET NUMBER: 480052.409C2
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 206-622-4900
  INFORMATION FOR SEQ ID NO: 2:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 162 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
    MOLECULE TYPE: protein
US-08-7626-817-2

```

Query Match

Query Match 17.3%; Score 98.5; DB 1; Length 162;
Best Local Similarity 26.1%; Pred. No. 0.00026;
Matches 30; Conservative 22; Mismatches 48; Indels 15; Gaps 5;

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	93.5	16.4	162	1	A53484	interleukin-15 pre
2	82	14.4	567	2	A29498	lymphocyte antigen
3	78.5	13.7	607	1	ABXL72	74K albumin precu
4	74	13.0	2197	2	B71600	variant-specific s
5	73.5	12.9	868	2	T31527	hypothetical prote
6	72.5	12.7	162	2	I49124	interleukin-15 - m
7	72.5	12.7	1101	2	T26919	hypothetical prote
8	72	12.6	262	2	F2858	probable methyl tr
9	72	12.6	262	2	T41813	ACMNPV orf69 - Bom
10	72	12.6	304	2	A32108	translation initia
11	71.5	12.5	195	2	S42022	ureidoglycolate hy
12	70.5	12.3	848	2	T00372	hypothetical prote
13	69	12.1	206	2	S49882	hypothetical prote
14	69	12.1	1220	2	T06403	resistance complex
15	68.5	12.0	336	2	S42632	Fit-15 protein pre
16	68.5	12.0	1257	2	T01020	hypothetical prote
17	68.5	12.0	1671	2	S71628	sensory transducti
18	68	11.9	329	2	C69483	hypothetical prote
19	68	11.9	397	2	A33880	syndecan 2 - human
20	67.5	11.8	406	2	T28957	hypothetical prote
21	67	11.7	2924	2	T18378	variant-specific s
22	66.5	11.6	420	2	A25876	vitellogenin III p
23	66.5	11.6	628	2	S61160	hypothetical prote
24	66.5	11.6	656	2	T37941	conserved hypother
25	66.5	11.6	1750	2	H64403	ribonucleoside-tri
26	66.5	11.6	3712	1	YGCEVC	alpha-aminoacidyl-
27	66	11.6	736	2	T00023	transcription fact
28	66	11.6	805	2	A46266	aryl hydrocarbon r
29	65.5	11.5	189	2	S11632	myosin regulatory

S29498
lymphocyte antigen Ly84 precursor - mouse
N:Alternate names: 38.5K T1 glycoprotein; ST2L protein
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Dec-2000
C:Accession: S29498; A33541; S17657; S07054
R:fangisawa, K.; Takagi, T.; Tsukamoto, T.; Tetsuka, T.; Tominaga, S.
FEBS Lett. 318, 83-87, 1993
A:Title: Presence of a novel primary response gene ST2L, encoding a product
A:Reference number: S29498; MUID: 93170492
A:Accession: S29498
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-567 <AND>
A:Cross-references: ENBL:13695; NID:q286100; PTDN:BAA02854.1; PID:q286101

R:Klemenz, R.; Hoffmann, S.; Werenskiold, A.K.
Proc. Natl. Acad. Sci. U.S.A. 86, 5708-5712, 1989
A:Title: Serum- and oncoprotein-mediated induction of a gene with sequence similarity to A33541; MUID:8934536
A:Reference number: A33541; MUID:8934536
A:Accession: A33541
A:Molecule type: mRNA
A:Residues: 1-191, 'V', 193-328, 'SKECPSHIA' <KLE>
A:Cross-references: GB:M24843; NID:9201103; PIDN:AAA40160.1; PID:g201104
R:Tominaga, S.I.; Jenkins, N.A.; Gilbert, D.J.; Copeland, N.G.; Tetsuka, T.
Biochim. Biophys. Acta 1090, 1-8, 1991
A:Title: Molecular cloning of the murine ST2 gene. Characterization and chromosomal mapping
A:Reference number: S17657; MUID:91355215
A:Accession: S17657
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-191, 'V', 193-328, 'SKECPSHIA' <TOM>
A:Cross-references: EMBL:X60184; NID:954200; PIDN:CAA42742.1; PID:g54201
R:Tominaga, S.I.
FEBS Lett. 258, 301-304, 1989
A:Title: A putative protein of a growth specific cDNA from BALB/c-3T3 cells is highly similar to the human interleukin-1 receptor type I
A:Reference number: S07034; MUID:90092495
A:Accession: S07034
A:Molecule type: mRNA
A:Residues: 1-328, 'SKECPSHIA' <TO2>
A:Cross-references: EMBL:Y07519; NID:g55517; PIDN:CAA68812.1; PID:g55518
A:Note: it is uncertain whether Met-1, Met-7 or Met-19 is the initiator
C:Genetics:
A:Gene: ST2
A:Map position: 1
A:Introns: 27/1; 97/2; 155/3; 210/1; 233/1; 280/2
A:Superfamily: interleukin-1 receptor type I
C:Keywords: glycoprotein
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-337/Product: ST2 protein #status predicted <MAT>
F:60.101,107,146,176,194,225,259,278/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 14.4%; Score 82; DB 2; Length 567;
Best Local Similarity 27.1%; Pred. No. 1.5;
Matches 29; Conservative 17; Mismatches 29; Indels 32; Gaps 6;
QY 7 DQKNVNDLVPEFLPAPDVETNCWFAFCFQAQLKASANTGNERNINVSIKLKRK 66
DB 114 PPSNIPDYLMTYRGSDKNFKITCTIDLYNTAPVQWFKNCAL 160
DB 76 DRK-----FLPARVE-----DSGIYACV-----IKSPND-NKGTGLNVTIHK---K 113
QY 67 PPSTN-----AGRRQKRLTCPCSDSYEKPPKEFLERFKSL 103
DB 114 PPSNIPDYLMTYRGSDKNFKITCTIDLYNTAPVQWFKNCAL 160
RESULT 3
ABXL72
74K albumin precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: B41682; S02693; A05288
R:Moakatis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.
Mol. Endocrinol. 3, 464-473, 1989
A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic acid during development.
A:Reference number: A41682; MUID:89313788
A:Accession: B41682
A:Molecule type: mRNA
A:Residues: 3-607 <MOS>
A:Cross-references: GB:M21442; NID:9213930; PIDN:AAA49637.1; PID:g213931
R:Schorpp, M.; Doebeiling, U.; Wagner, U.; Ryffel, G.U.
J. Mol. Biol. 199, 83-93, 1988
A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Deletion of the 5'-proximal exon results in a nontranslatable mRNA
A:Reference number: S02692; MUID:88172470
A:Accession: S02693
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-48 <SCH>

R:Klemenz, R.; Hoffmann, S.; Werenskiold, A.K.
Proc. Natl. Acad. Sci. U.S.A. 86, 5708-5712, 1989
A:Title: Serum- and oncoprotein-mediated induction of a gene with sequence similarity to A33541; MUID:8934536
A:Reference number: A33541; MUID:8934536
A:Accession: A33541
A:Molecule type: mRNA
A:Residues: 1-191, 'V', 193-328, 'SKECPSHIA' <KLE>
A:Cross-references: GB:M24843; NID:9201103; PIDN:AAA40160.1; PID:g201104
R:Tominaga, S.I.; Jenkins, N.A.; Gilbert, D.J.; Copeland, N.G.; Tetsuka, T.
Biochim. Biophys. Acta 1090, 1-8, 1991
A:Title: Molecular cloning of the murine ST2 gene. Characterization and chromosomal mapping
A:Reference number: S17657; MUID:91355215
A:Accession: S17657
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-191, 'V', 193-328, 'SKECPSHIA' <TOM>
A:Cross-references: EMBL:X60184; NID:954200; PIDN:CAA42742.1; PID:g54201
R:Tominaga, S.I.
FEBS Lett. 258, 301-304, 1989
A:Title: A putative protein of a growth specific cDNA from BALB/c-3T3 cells is highly similar to the human interleukin-1 receptor type I
A:Reference number: S07034; MUID:90092495
A:Accession: S07034
A:Molecule type: mRNA
A:Residues: 1-328, 'SKECPSHIA' <TO2>
A:Cross-references: EMBL:Y07519; NID:g55517; PIDN:CAA68812.1; PID:g55518
A:Note: it is uncertain whether Met-1, Met-7 or Met-19 is the initiator
C:Genetics:
A:Gene: ST2
A:Map position: 1
A:Introns: 27/1; 97/2; 155/3; 210/1; 233/1; 280/2
A:Superfamily: interleukin-1 receptor type I
C:Keywords: glycoprotein
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-337/Product: ST2 protein #status predicted <MAT>
F:60.101,107,146,176,194,225,259,278/Binding site: carbohydrate (Asn) (covalent) #status predicted

A:Cross-references: EMBL:226826
R:Wolfe, A.P.; Glover, J.F.; Martin, S.C.; Tenniswood, M.P.R.; Williams, J.L.; Tata, E.R. J. Biochem. 146, 489-496, 1985
A:Title: Deinduction of transcription of xenopus 74-kDa albumin genes and destabiliza
A:Reference number: A05288; MUID:85126974
A:Accession: A05288
A:Molecule type: mRNA
A:Residues: 459-502, 'L', 504-557 <WOL>
A:Cross-references: GB:M28276
A:Note: the authors translated the codon TAT for residue 63 as Thr
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds coppe
mones (weak bonds with these hormones promote their transfer across the membranes), t
C:Genetics: 27/1
A:Introns: 27/1
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; glycoprotein; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: 74K serum albumin #status predicted <MAT>
F:32-201/Domain: serum albumin repeat homology <SA1>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:30/Binding site: copper (His) #status predicted
F:80-88,101-117,116-127,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-3
F:256/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.7%; Score 78.5; DB 1; Length 607;
Best Local Similarity 24.6%; Pred. No. 3.6;
Matches 35; Conservative 18; Mismatches 48; Indels 41; Gaps 5;
QY 4 DIVDQLKNVNDLVPEF-----LPAPDVETNCWFAFCFQAQLKASAN 48
DB 72 EINDFAKSCINDTPECEKPVGTLLFDKLCADPAVGNYEWSKCCAKQDPERAQCFFAH 131
QY 49 TGNNERII----NVSIKLKRK-----PPSTNAGRRQKRLTCPS 84
DB 132 RDHEHTSIKPEPETCKLKEHPDLDLSAFTHFARNHPDLYPPAVLALTKYHKLAFHC 191
QY 85 CDSYERKPPKEFLERFKSLLOK 106
DB 192 CEEDKE--KCFSEKKQLMKQ 211
RESULT 4
B71600
variant-specific surface protein 1 homolog PFB1055c - malaria parasite (Plasmodium fa
N:Alternate names: erythrocyte membrane binding protein 1 (EMPI)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C:Accession: B71600
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.
; Perlea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743
A:Accession: B71600
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2197 <GAR>
A:Cross-references: GB:AE001434; GB:AE001362; NID:g3845341; PIDN:AACT71996.1; PID:g384
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB1055c
Query Match 13.0%; Score 74; DB 2; Length 2197;
Best Local Similarity 29.5%; Pred. No. 37;
Matches 28; Conservative 16; Mismatches 41; Indels 10; Gaps 4;
QY 9 LKNVNDLVPEFLPAPDVETNCWFAFCFQAQLKASANTGNERNINVSIKLKRK 66
DB 1594 KSFLETWIPK-IADVNDQDNVTKLSKFGSCGCSAISTN-GNEEDADICMIKKLEKK 1651

T26919
hypothetical protein Y45F10B.10 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26919
R:McMurray, A.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z20286
A:Accession: T26919
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1101 <WIL>
A:Cross-references: EMBL:AL021487; PTDN:CAA16357.1; GSPDB:GN00022; CESP:Y45F10B.10
A:Experimental source: Clone Y45F10B
C:Genetics:
A:Gene: CESP:Y45F10B.10
A:Map position: 4
A:Introns: 217/3; 292/1; 451/3; 556/2; 637/3; 755/3; 823/2; 946/1; 1078/3

```

Query Match      12.7%; Score 72.5; DB 2; Length 1101;
Best Local Similarity 31.5%; Pred. No. 26;
Matches 23; Conservative 14; Mismatches 23; Indels 13; Gaps 4;

Qy 40 QKAQLKSANTGNERRIINVSITKKLRKP--PSTNAGRQK-HRUTCPSCDSYEKKPKPEF 96
      : : : : : : : : : : : : : : : : : : : : : :
Db 1013 RSRAQSVSSASNEPVASTSAGTEIKKDPILSSNNGNAQSPRATAP-----KPTDM 1065

Qy 97 LERFK---SLLQK 106
      ||| | ||::|
Db 1066 LERSKSRSTSLIEK 1078

```

RESULT 8
F72858
probable methyl transferase - Autographa californica nuclear polyhedrosis virus
C:Species: Autographa californica nuclear polyhedrosis virus, ACNMPV
A:Note: dsDNA virus
C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 12-Nov-1999
C:Accession: F72858
R:Viroyes, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.
Virology 202, 586-605, 1994
A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis vir
A:Reference number: A72850; MUID:94303173
A:Accession: F72858
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <AYR>
A:Cross-references: GB:L22858; NID:G510708; PIDN:AAA66699.1; PID:G559138
C:Genetics:
A:Gene: ACOrf-69

```

Query Match          12.6%; Score 72; DB 2; Length 262;
Best Local Similarity 21.7%; Pred. No. 6.8;
Matches 18; Conservative 20; Mismatches 35; Indels 10; Gaps 12;

QY 30 NCEWSAFSCFOKAOLKSANTGNERRIINVSITKLLKRPSPSTNAGRQKH-----RLTCPS 84
    || | | : : | : : | : | : | : : | : | : | : | : | : | : | : |
Db 176 NCVLKVFDAFEHETIOMLN-----KFNHFHFKWLYKRPSPSPANSERYLCIFKNLVRPY 230
    || | | : : | : : | : | : | : : | : | : | : | : | : | : | : |

QY 85 CDSYEKKPPKEFLERFKSLLOQM 107
    || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 231 CNNYVNELEKQFKYYRIQLKNL 253
    || : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 9
T41813
ACMNPVorf6 - Bombyx mori nuclear polyhedrosis virus (isolate T3)
C:Species: Bombyx mori nuclear polyhedrosis virus, BmSNPV
A:Variety: isolate T3
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C:Accession: T41813

```

R:Gomi, S.; Majima, K.; Maeda, S.
J. Gen. Virol. 80, 1323-1337, 1999
A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
A:Reference number: Z22020; MUID:99281911
A:Accession: T41813
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-262 <KAM>
A:Cross-references: EMBL:L33180; PIDN:AAC63742.1
A:Experimental source: isolate T3
C:Genetics:
A:Note: Orf_57

Query Match 12.6%; Score 72; DB 2; Length 262;
Best Local Similarity 21.7%; Pred. No. 6.8;
Matches 18; Conservative 19; Mismatches 36; Indels 10; Gaps 2;
2;
Qy 30 NCEWSAFCTQKAOLKSANTGNRIINVSILKLRKPPSTNAGRQKH-----RLTCPS 84
Db 176 NCVLKVFDAFEHKTQMLN-----KFVNHFEKWLKPPSSRPANSERYLICFNKLVRPY 230
Qy 85 CDSYEKRPKPEFLERFKSLQKM 107
Db 231 CNDYVNELEKQFKYYRIOLKNL 253

RESULT 10
A32108
translation initiation factor eIF-2 alpha chain - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein J1429; protein YJR007W
C:Species: Saccharomyces cerevisiae
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 26-Aug-1999
A:Accession: A32108; S55195; S57022
R:Cigan, A.M.; Pabich, E.K.; Feng, L.; Donahue, T.F.
Proc. Natl. Acad. Sci. U.S.A. 86, 2784-2788, 1989
A:Title: Yeast translation initiation suppressor su12 encodes the alpha-subunit of eukaryotic translation initiation factor eIF-2
A:Reference number: A32108; MUID:8920411
A:Accession: A32108
A:Molecule type: DNA
A:Residues: 1-304 <CIG>
A:Cross-references: EMBL:M25552; NID:g341369; PIDN:AA70332.1; PID:g903889
R:de Haan, M.; Smits, P.H.M.; Grivell, L.A.; Smits, P.H.M.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55183
A:Accession: S55195
A:Molecule type: DNA
A:Residues: 1-304 <DEH>
A:Cross-references: EMBL:X87611; NID:g854567; PIDN:CAA60929.1; PID:g854580
R:de Haan, M.; Grivell, L.A.; Smits, P.H.M.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56771
A:Accession: S57022
A:Molecule type: DNA
A:Residues: 1-304 <ZAG>
A:Cross-references: EMBL:Z49507; NID:g1015631; PIDN:CAA89529.1; PID:g1015632; MIPS:YJR007W
C:Genetics:
A:Gene: SGD:SUI2
A:Cross-references: SGD:S0003767; MIPS:YJR007W
A:Map position: 10R
C:Superfamily: translation initiation factor eIF-2 alpha chain
C:Keywords: phosphoprotein; protein biosynthesis

Query Match 12.6%; Score 72; DB 2; Length 304;
Best Local Similarity 33.3%; Pred. No. 7.9;
Matches 20; Conservative 14; Mismatches 22; Indels 4; Gaps 4;
4
Qy 4 DIVDLKYNVD-LVPEFLPAPEDVETNC-EWSAFSCFQKALKSANTGNRIINVSILK 61
Db 164 DVLDLKNYISKRLTPQAVKTRADVESCFSYEGIDAIDKA-LKSAEDMSTEQ-MQVRVK 221

R:Gomi, S.; Majima, K.; Maeda, S.
J. Gen. Virol. 80, 1323-1337, 1999
A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
A:Reference number: Z22020; MUID:99281911
A:Accession: T41813
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-262 <KAM>
A:Cross-references: EMBL:L33180; PIDN:AAC63742.1
A:Experimental source: isolate T3
C:Genetics:
A:Note: Orf_57

Query Match 12.6%; Score 72; DB 2; Length 262;
Best Local Similarity 21.7%; Pred. No. 6.8;
Matches 18; Conservative 19; Mismatches 36; Indels 10; Gaps 2;
2;
Qy 30 NCEWSAFCTQKAOLKSANTGNRIINVSILKLRKPPSTNAGRQKH-----RLTCPS 84
Db 176 NCVLKVFDAFEHKTQMLN-----KFVNHFEKWLKPPSSRPANSERYLICFNKLVRPY 230
Qy 85 CDSYEKRPKPEFLERFKSLQKM 107
Db 231 CNDYVNELEKQFKYYRIOLKNL 253

RESULT 10
A32108
translation initiation factor eIF-2 alpha chain - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein J1429; protein YJR007W
C:Species: Saccharomyces cerevisiae
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 26-Aug-1999
A:Accession: A32108; S55195; S57022
R:Cigan, A.M.; Pabich, E.K.; Feng, L.; Donahue, T.F.
Proc. Natl. Acad. Sci. U.S.A. 86, 2784-2788, 1989
A:Title: Yeast translation initiation suppressor su12 encodes the alpha-subunit of eukaryotic translation initiation factor eIF-2
A:Reference number: A32108; MUID:8920411
A:Accession: A32108
A:Molecule type: DNA
A:Residues: 1-304 <CIG>
A:Cross-references: EMBL:M25552; NID:g341369; PIDN:AA70332.1; PID:g903889
R:de Haan, M.; Smits, P.H.M.; Grivell, L.A.; Smits, P.H.M.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55183
A:Accession: S55195
A:Molecule type: DNA
A:Residues: 1-304 <DEH>
A:Cross-references: EMBL:X87611; NID:g854567; PIDN:CAA60929.1; PID:g854580
R:de Haan, M.; Grivell, L.A.; Smits, P.H.M.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56771
A:Accession: S57022
A:Molecule type: DNA
A:Residues: 1-304 <ZAG>
A:Cross-references: EMBL:Z49507; NID:g1015631; PIDN:CAA89529.1; PID:g1015632; MIPS:YJR007W
C:Genetics:
A:Gene: SGD:SUI2
A:Cross-references: SGD:S0003767; MIPS:YJR007W
A:Map position: 10R
C:Superfamily: translation initiation factor eIF-2 alpha chain
C:Keywords: phosphoprotein; protein biosynthesis

Query Match 12.6%; Score 72; DB 2; Length 304;
Best Local Similarity 33.3%; Pred. No. 7.9;
Matches 20; Conservative 14; Mismatches 22; Indels 4; Gaps 4;
4
Qy 4 DIVDLKYNVD-LVPEFLPAPEDVETNC-EWSAFSCFQKALKSANTGNRIINVSILK 61
Db 164 DVLDLKNYISKRLTPQAVKTRADVESCFSYEGIDAIDKA-LKSAEDMSTEQ-MQVRVK 221

RESULT 11
S42022
ureidoglycolate hydrolase (EC 3.5.3.19) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YJR032C
C:Species: Saccharomyces cerevisiae
C:Date: 31-Mar-1992 #sequence_revision 14-Sep-1994 #text_change 21-Jul-2000
A:Accession: S42022; S48494
R:Yoo, H.S.; Cooper, T.G.
Yeast 7, 693-698, 1991
A:Title: The ureidoglycollate hydrolase (DAL3) gene in Saccharomyces cerevisiae.
A:Reference number: S42022; MUID:92133160
A:Accession: S42022
A:Molecule type: DNA
A:Residues: 1-195 <YOO>
A:Cross-references: EMBL:M64778; NID:g171369; PIDN:AAA73025.1; PID:g171370
R:Rowley, K.
submitted to the EMBL Data Library, October 1994
A:Reference number: S48478
A:Accession: S48494
A:Molecule type: DNA
A:Residues: 1-195 <ROW>
A:Cross-references: GB:Z47047; EMBL:Z38061; NID:g603997; PID:g7633377; MIPS:YJR032C
C:Genetics:
A:Gene: SGD:DAL3
A:Cross-references: SGD:S0001471; MIPS:YJR032C
A:Map position: 9R
C:Keywords: allantoin degradation; hydrolase; lipoprotein; methylated carboxyl end; p
F.192/Binding site: farnesyl (Cys) (covalent) #status predicted
F.192/Modified site: methyl ester carboxyl end (Cys) (in mature form) #status predict

Query Match 12.5%; Score 71.5; DB 2; Length 195;
Best Local Similarity 29.2%; Pred. No. 5.6;
Matches 19; Conservative 9; Mismatches 24; Indels 13; Gaps 1;
Qy 6 VDQLKYNVDLYPEFLPAPEDVETNC-EWSAFSCFQKALKSANTGNRIINVSILKLR 65
Db 48 VSQVENKSTKVP-----NWLFRCPQPHLNRVFTQGSNQAISHIKVLEK 94

Qy 66 KPPST 70
Db 95 HPCST 99

RESULT 12
T00372
hypothetical protein KIAA0650 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 17-Nov-2000
A:Accession: T00372; T12523
R:Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete coding sequences of 1000 human genes were determined by cDNA sequencing and
A:Reference number: T12523
A:Accession: T00372
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-848 <ISH>
A:Cross-references: EMBL:AB014550; NID:g3327113; PIDN:BAA31625.1; PID:g3327114
A:Experimental source: brain
R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z17524
A:Accession: T12523
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-848 <WAM>
A:Cross-references: EMBL:AL080138
A:Experimental source: adult testis; clone DKFZp434K063
C:Genetics:
A:Note: KIAA0650; DKFZp434K063.1

Query Match 12.3%; Score 70.5; DB 2; Length 848;
Best Local Similarity 25.9%; Pred. No. 32;
Matches 21; Conservative 17; Mismatches 38; Indels 5; Gaps 3;
QY 19 EFLPAPEDVTNC-ENSAFSCFOKAOLKSANTGNRIINVSIIKKLRKPPSTNAGRQK 77
Db 66 KFTPGPGNKDLCFTWRESDFIRVOLIS---GPPAKLLIDNPDELKESIPVIN-GRDLQ 121
QY 78 HRLTSCDSYKPKPEFLE 98
Db 122 NPIIVQLCDQWDNPAPVQHVK 142
RESULT 13
S49882
hypothetical protein Y1127c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein Y18277.O2c
C:Species: Saccharomyces cerevisiae
C:Date: 13-Jan-1995 #sequence_revision 20-Feb-1995 #text_change 05-Nov-1999
C:Accession: S49882
R:Hamlyn, N.; Churcher, C.
submitted to the EMBL Data Library, November 1994
A:Reference number: S49881
A:Accession: S49882
A:Molecule type: DNA
A:Residues: 1-206 <HAM>
A:Cross-references: GB:Z47047; EMBL:Z46833; NID:g603997; PID:g763219; GSPDB:GN00009; MIP
C:Genetics:
A:Gene: MIPS:Y1127C
A:Map position: 9L
Query Match 12.1%; Score 69; DB 2; Length 206;
Best Local Similarity 26.0%; Pred. No. 11;
Matches 19; Conservative 14; Mismatches 40; Indels 0; Gaps 0;
QY 8 QLNKYNDLVPEFLPAPEDVTNCWSAFSCFOKAOLKSANTGNRIINVSIIKKLRKPP 67
Db 12 QATSVVNGLLSNLPGVPKIRANNKTSVNGSKAQLIDRLNKKRVQLQNRDVHIKKKC 71
QY 68 PSTNAGRQKHRL 80
Db 72 KLVKKKKVKKHKL 84
RESULT 14
T06403
resistance complex protein I2C-1 - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 17-Nov-2000
C:Accession: T06403
R:Ori, N.; Eshed, Y.; Paran, I.; Presting, G.; Aviv, D.; Tanksley, S.; Zamir, D.; Fluhr,
Plant Cell 9, 521-532, 1997
A:Title: The I2C family from the wilt disease resistance locus I2 belongs to the nucleot
A:Reference number: Z15652; MUID:97290204
A:Accession: T06403
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1220 <ORI>
A:Cross-references: EMBL:AF004878; NID:g2258314; PIDN:AAB63274.1; PID:g2258315
C:Genetics:
A:Gene: I2C-1
A:Map position: 11
C:Function:
A:Description: confers resistance against Fusarium oxysporum
C:Superfamily: disease resistance protein RPS2; leucine-rich alpha-2-glycoprotein repeat
Query Match 12.1%; Score 69; DB 2; Length 1220;
Best Local Similarity 30.3%; Pred. No. 65;
Matches 30; Conservative 8; Mismatches 23; Indels 38; Gaps 6;
QY 13 VNDLVPEFLPA-----PEDVET-----NCEWSAFSCFOKAOLKSANTGN 51

Db 948 VDDISPELPATROLSIENCHNVTRFLIPTATESLHIRNCEKLSMACGGAOLTSLNW- 1006
QY 52 NERIINVSIIKKLRKP---PSTNAGRQKHRLT-CPSCD 86
Db 1007 -----GCKKLKCLPELLPSL-----KELRLTYCPEIE 1033
RESULT 15
S42632
Fit-1S protein precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 25-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 19-May-2000
C:Accession: S42632
R:Bergers, G.; Reikerstorfer, A.; Braselmann, S.; Graninger, P.; Busslinger, M.
EMBO J. 13, 1176-1188, 1994
A:Title: Alternative promoter usage of the Fos-responsive gene Fit-1 generates mRNA 1
A:Reference number: S42632; MUID:94178260
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-336 <BER>
A:Cross-references: GB:U04319; NID:g488278; PIDN:AAA67172.1; PID:g488279
C:Superfamily: vaccinia virus B15R protein; immunoglobulin homology
Query Match 12.0%; Score 68.5; DB 2; Length 336;
Best Local Similarity 22.4%; Pred. No. 20;
Matches 28; Conservative 25; Mismatches 35; Indels 37; Gaps 6;
QY 3 IDIVDLKKNVNDLVP-----EFLPAPEDVTNCWSAFSCFOKAOLKSAN 48
Db 49 INPVEWYYSNTNRIPTQKRNRIFVSRDLKFLPAKVE-----DSGIYTCVIRSP-ESTK 102
QY 49 TGNNERIINVSIIKKLRKPP-----STNAGRQKHRLTSCDSYKPKPEFLE 98
Db 103 TGS-----LNVITYK---RPPNCKIPDYMYSTVDGSKNSKITCPTIALYNNWTAPVQWFK 155
QY 99 RFKSL 103
Db 156 NCKAL 160
Search completed: May 23, 2001, 11:12:48
Job time: 251 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:11:42 ; Search time 40.06 seconds
(without alignments)
92.351 Million cell updates/sec

Title: US-09-522-217-2_COPY_41_148

Perfect score: 571

Sequence: 1 QDIVDOLKNVNDLPEF.....EKKPKPELFERKSLQKMI 108

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100.5	17.6	162	1 IL15_FELCA	O97687 felis silve
2	98.5	17.3	162	1 IL15_HUMAN	P40933 homo sapien
3	93.5	16.4	162	1 IL15_CERAE	P40221 cercopitheci
4	93.5	16.4	162	1 IL15_MACMU	P48092 macaca mula
5	89.5	15.7	162	1 IL15_BOVIN	Q28028 bos taurus
6	82	14.4	337	1 IRL1_MOUSE	P14719 mus musculu
7	79.5	13.9	162	1 IL15_PIG	Q95253 sus scrofa
8	78.5	13.7	607	1 ALB2_XENLA	P14872 xenopus lae
9	72.5	12.7	162	1 IL15_MOUSE	P48346 mus musculu
10	72	12.6	262	1 Y069_NPVAC	P41469 autoqrpha
11	72	12.6	304	1 IF2A_YEAST	P20459 saccharomyc
12	71.5	12.5	195	1 ALIA_YEAST	P32459 saccharomyc
13	69	12.1	206	1 YIM7_YEAST	P40470 saccharomyc
14	69	12.1	741	1 IDH_AZOVI	P16100 azotobacter
15	68	11.9	201	1 SDC2_HUMAN	P34741 homo sapien
16	67.5	11.8	162	1 IL15_RAT	P97604 rattus norv
17	67.5	11.8	300	1 SPY4_MOUSE	O9wtb2 mus musculu
18	66.5	11.6	420	1 VIT3_DROME	P06607 drosophila
19	66.5	11.6	1750	1 Y832_METJA	Q58242 methanococc
20	66.5	11.6	3712	1 ACVS_CEPAC	P25464 cephalospor
21	66	11.6	736	1 BACL_HUMAN	O14867 homo sapien
22	66	11.6	805	1 AHR_MOUSE	P30561 mus musculu
23	65.5	11.5	576	1 P80C_HUMAN	P38432 homo sapien
24	65	11.4	695	1 TKT_PICST	P34736 pichia stip
25	65	11.4	739	1 BACL_MOUSE	P97302 mus musculu
26	64.5	11.3	484	1 CCB3_MOUSE	P54285 mus musculu
27	64.5	11.3	4385	1 YP73_CAEEL	Q09222 caenorhabdi
28	64	11.2	921	1 SYI_BACSU	Q45477 bacillus su
29	63.5	11.1	173	1 YNN3_YEAST	P53913 saccharomyc
30	63.5	11.1	457	1 ODR7_CAEEL	P41933 caenorhabdi
31	63.5	11.1	477	1 CCB3_RABIT	P54286 oryctolagus
32	63.5	11.1	484	1 CCB3_HUMAN	P54284 homo sapien
33	63.5	11.1	484	1 CCB3_RAT	P54287 rattus norv

34	63.5	11.1	525	1 UL32_EBV	P03184 epstein-bar
35	63.5	11.1	789	1 YKZ5_CAEEL	P34332 caenorhabdi
36	63.5	11.1	1636	1 BUD3_YEAST	P25558 saccharomyc
37	63.5	11.1	2476	1 ATRX_MOUSE	Q61687 mus musculu
38	63	11.0	403	1 MTB1_BREPP	P10283 brevibacter
39	63	11.0	525	1 KBN8_YEAST	P38070 saccharomyc
40	63	11.0	864	1 YG3M_YEAST	P48237 saccharomyc
41	63	11.0	1206	1 FM14_MOUSE	Q03859 mus musculu
42	63	11.0	2375	1 ATRX_HUMAN	P46100 homo sapien
43	62.5	10.9	1468	1 FMN1_MOUSE	Q03860 mus musculu
44	62.5	10.9	2210	1 RRPQ_TACV	P20430 tacaribe vi
45	62	10.9	345	1 AST4_DROME	P10084 drosophila

ALIGNMENTS

RESULT 1	
IL15_FELCA	
ID IL15_FELCA	STANDARD; PRT; 162 AA.
AC O97687;	
DT 30-MAY-2000 (Rel. 39, Created)	
DT 30-MAY-2000 (Rel. 39, Last sequence update)	
DT 30-MAY-2000 (Rel. 39, Last annotation update)	
DE INTERLEUKIN-15 PRECURSOR (IL-15).	
GN IL15.	
OS Felis silvestris catus (Cat).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.	
OX NCBI_TaxID=9685;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC TISSUE=Lymph node;	
RA Barger A.B., Dean G.A., Lavoy A.S.;	
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.	
CC FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-	
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15	
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R	
CC GAMMA BUT NOT IL-2R ALPHA (BY SIMILARITY).	
CC SUBCELLULAR LOCATION: SECRETED	
CC SIMILARITY: BELONGS TO THE IL-15 FAMILY.	
CC	
CC This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC the European Bioinformatics Institute. There are no restrictions on its	
CC use by non-profit institutions as long as its content is in no way	
CC modified and this statement is not removed. Usage by and for commercial	
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC or send an email to license@isb-sib.ch).	
CC	
DR EMBL: AF108148; AAD05268.1; --	
KW Cytokine; Glycoprotein; Signal.	
FT SIGNAL	1 29 POTENTIAL.
FT PROPEP	30 48 POTENTIAL.
FT CHAIN	49 162 INTERLEUKIN-15.
FT DISULFID	83 133 POTENTIAL.
FT DISULFID	90 136 POTENTIAL.
FT CARBOHYD	104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE	162 AA; 18412 MW; DBC7CF7F40110DD CRC64;

Query Match	17.6%	Score	100.5;	DB 1;	Length	162;			
Best Local Similarity	26.7%;	Pred. No.	0.003;						
Matches	31;	Conservative	18;	Mismatches	48;	Indels	19;	Gaps	5;
Qy	4	DIVDOLK---	NYVNDL-VPEFLPAPEDVETNCWEAFSCF----	OKAOLKSANTGNERNI	55				
		:	:	:	:	:	:	:	:
Db	52	DVSDLKIKIDKIQLSHDITATLYTESDHPNCKVTKMCKFLLLEHLHVLSKSNETHIQIV	111						
Qy	56	INVSITKLKRPSPNAGRQKRLT---	CPSCDSYEKKPKPELFERKSLQKMI	108					
		:	:	:	:	:	:	:	:
Db	112	ENIIT-----	LANSGLSSNRNITETGCKECEELEKNEIKFQLQSFVHIQVMEI	159					


```
CC -----
DR EMBL; U03099; AA18416.1; ..
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 127 127 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 162 AA; 18222 MW; 1BF9A82644E1C9B7 CRC64;

Query Match 16.4%; Score 93.5; DB 1; Length 162;
Best Local Similarity 26.1%; Pred. No. 0.015;
Matches 30; Conservative 22; Mismatches 48; Indels 15; Gaps 5;

QY 3 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKAQKKSANTGNNE 53
DB 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVISHESGDTDIHD 109
QY 54 RIINVSIKKLRKPPSTNAGRQKRLTCPCSDSYEKPKPPKEFLERFKSLQKMI 108
DB 110 TVENLII--LANNILSSNGNITES---GCKECELEEKNIKEFLQSFVHIVQMF 159

RESULT 4
IL15_MACMU
ID IL15_MACMU STANDARD; PRT; 162 AA.
AC P48092;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=96003435; PubMed=7561102;
RA Villinger F.J., Brar S.S., Mayne A.E., Chikkala N., Ansari A.A.;
RT "Comparative sequence analysis of cytokine genes from human and
RL nonhuman primates."
RL J. Immunol. 155:3946-3954(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Tatsumi M.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U19843; AAB60398.1; ..
DR EMBL; AB000555; BAA19149.1; ..
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133
SQ SEQUENCE 162 AA; 18554 MW; 6A63CAA329EB8302 CRC64;
```

Query Match 15.7%; Score 89.5; DB 1; Length 162;
Best Local Similarity 25.7%; Pred. No. 0.036;

```
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 127 127 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARIANT 5 5 K -> T.
FT VARIANT 31 31 I -> T.
SQ SEQUENCE 162 AA; 18194 MW; D233CF7FF6188C01 CRC64;

Query Match 16.4%; Score 93.5; DB 1; Length 162;
Best Local Similarity 26.1%; Pred. No. 0.015;
Matches 30; Conservative 22; Mismatches 48; Indels 15; Gaps 5;

QY 3 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKAQKKSANTGNNE 53
DB 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVISHESGDTDIHD 109
QY 54 RIINVSIKKLRKPPSTNAGRQKRLTCPCSDSYEKPKPPKEFLERFKSLQKMI 108
DB 110 TVENLII--LANNILSSNGNITES---GCKECELEEKNIKEFLQSFVHIVQMF 159

RESULT 5
IL15_BOVIN
ID IL15_BOVIN STANDARD; PRT; 162 AA.
AC Q28028;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HOLSTEIN;
RX MEDLINE=97426124; PubMed=9282828;
RA Canals A., Gasbarre L.C., Boyd P.C., Almeria S., Zarlenga D.S.;
RT "Cloning and expression of bovine interleukin-15: analysis and
RT modulation of transcription by exogenous stimulation."
RL J. Interferon Cytokine Res. 17:473-480(1997).
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U42433; AAA85130.1; ..
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 104 104 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 113 113 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 121 121 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 127 127 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 162 AA; 18554 MW; 6A63CAA329EB8302 CRC64;
```


Matches	28;	Conservative	16;	Mismatches	58;	Indels	7;	Gaps
Oy	5	IVDOLKNVNDLVPEF-----LPAPEDVTNCWSAFSCFOKAQLKSANTGNNGRIINVS	59	:	: :	:	:	:
Dd	53	VINDLKT-IEHLIQSIHMDATLYTESDAHFNCKVKYTAQCFLLELRVLHESKNAIYEI-	110	:	:	:	:	:
Oy	60	IKKLKRPPSTNAGRRQKHRLTCPSDCSDSYEKKKPKPFERFKSLLOXMI	108	:	:	:	:	:
Dd	111	IENLTMLANLNLSIENTKGCEKEELESEKIIFLKFSFHVIQMFI	159	:	:	:	:	:
<p style="text-align:center;">RESULT 6</p>								
ID	IRLI_MOUSE	STANDARD; PRT:	337 AA.					
AC	P14719;							
DT	01-APR-1990 (Rel. 14, Created)							
DT	01-APR-1990 (Rel. 14, Last sequence update)							
DE	01-OCT-2000 (Rel. 40, Last annotation update)							
DE	INTERLEUKIN 1 RECEPTOR-LIKE 1 PRECURSOR (ST2 PROTEIN) (T1 PROTEIN)							
DE	(LYMPHOCYTE ANTIGEN 84).							
GS	ILIRLI OR ST2 OR STE2 OR LY84.							
OS	Mus musculus (Mouse).							
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.							
OX	NCBI_TaxID=10090;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RC	STRAIN=BALB/C;							
RX	MEDLINE=90092495; PubMed=2532153;							
RA	Tominaga S.;							
RT	"A putative protein of a growth specific cDNA from BALB/c-3T3 cells							
RT	is highly similar to the extracellular portion of mouse interleukin 1							
RT	receptor.";							
RL	FEBIS Lett. 258:301-304(1989).							
RN	[2]							
RP	SEQUENCE FROM N.A.							
RC	STRAIN=C3H/HE; TISSUE=Spleen;							
RX	MEDLINE=91355215; PubMed=1832015;							
RA	Tominaga S.I., Jenkins N.A., Gilbert D.J., Copeland N.G.,							
RA	Tetsuka T.;							
RT	"Molecular cloning of the murine ST2 gene. Characterization and							
RT	chromosomal mapping.";							
RL	Biochim. Biophys. Acta 1090:1-8(1991).							
RN	[3]							
RP	SEQUENCE FROM N.A.							
RX	MEDLINE=89345536; PubMed=2527364;							
RA	Klemenz R., Hoffmann S., Werenskiold A.K.;							
RT	"Serum- and oncprotein-mediated induction of a gene with sequence							
RT	similarity to the gene encoding carcinoembryonic antigen.";							
RL	Proc. Natl. Acad. Sci. U.S.A. 86:5708-5712(1989).							
CC	- FUNCTION: POSSIBLY INVOLVED IN REGULATION OF T-LYMPHOCYTE							
CC	ACTION.							
CC	- DEVELOPMENTAL STAGE: GROWTH-SPECIFIC EXPRESSION, GI-PHASE OF							
CC	CELL CYCLE.							
CC	- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.							
CC	- SIMILARITY: STRONG, TO INTERLEUKIN-1 RECEPTORS.							
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration							
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation							
CC	the European Bioinformatics Institute. There are no restrictions on its							
CC	use by non-profit institutions as long as its content is in no way							
CC	modified and this statement is not removed. Usage by and for commercial							
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/							
CC	or send an email to license@lsb-sib.ch).							
DR	EMBL; Y07519; CAA68812.1; -							
DR	EMBL; X60184; CAAA2742.1; -							
DR	EMBL; M24843; AAA40160.1; -							
DR	PIR; S07054; S07054							
DR	MGI; MG1.98427; Lyr84.							
DR	InterPro; IPR003006; -							
DR	Pfam; PF00047; Ig; 3.							
DR	Immunoglobulin domain; Glycoprotein; Signal.							

FT	SIGNAL	1	26	POTENTIAL.
FT	CHAIN	27	337	INTERLEUKIN 1 RECEPTOR-LIKE 1.
FT	DOMAIN	35	100	IG-LIKE C2-TYPE DOMAIN.
FT	DOMAIN	132	194	IG-LIKE C2-TYPE DOMAIN.
FT	DOMAIN	233	315	IG-LIKE C2-TYPE DOMAIN.
FT	DISULFID	42	93	BY SIMILARITY.
FT	DISULFID	139	187	BY SIMILARITY.
FT	DISULFID	240	308	BY SIMILARITY.
FT	CARBOHYD	60	60	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	101	101	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	107	107	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	146	146	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	176	176	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	194	194	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	225	225	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	259	259	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	278	278	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	VARIANT	192	192	A -> V (IN STRAIN C3H/HE).
SO	SEQUENCE	337 AA;	38502 MW;	7574372722486926 CRC64;

Query Match	14.48;	Score 82;	DB 1;	Length 337;
Best Local Similarity	27.18;	Pred. No. 0.45;		
Matches 28;	Conservative	17;	Mismatches 29;	Indels 32; Gaps 6;

Qy	7	DQLKNYNDVLPELPAPEDVETNCW	SFAFCQKALKSANTGNNERINVS	IKLKRK 66
			: : : :	: : : : : :
Db	76	DRLK-----FLPARVE-----	DSGIYACV----	IRSPNL-NKGYLNVTHK---K 113
			: : : :	: : : : : :
Qy	67	PPSTN-----AGRRQKHRLTCPS	CDSCSYEKKPKPEFLERFKSL	103
			: : : : : :	: : : : : :
Db	114	PSCNIPDYLMYSTVRGSDKNFKIT	CPIDLYNWTAPVQWFKNCKAL	160

RESULT	7
IL15_PIG	
ID	IL15_PIG
AC	Q95253;
CD	01-NOV-1997 (Rel. 35, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)
DT	15-DEC-1998 (Rel. 37, Last annotation update)
DE	INTERLEUKIN-15 PRECURSOR (IL-15).
GN	IL15.
OS	Sus scrofa (Pig).
OC	Eukaryota; Metazoa;
OC	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX	NCBI_TaxID=9623;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Blood;
RX	MEDLINE=97449311; PubMed=9305780;
RA	Canals A., Grimm D.R., Gasbarre L.C., Lunney J.K., Zarlenga D.S.;
RT	"Molecular cloning of cDNA encoding porcine Interleukin-15.;"
RL	Gene 195:337-339(1997).
CC	-!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC	LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC	WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC	GAMMA BUT NOT IL-2R ALPHA (BY SIMILARITY).
CC	SUBCELLULAR LOCATION: SECRETED.
CC	-!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC	or send an email to license@isb-sib.ch).
CC	-----
EMBL;	U58142; AAB72031.1; -
DR	Cytokine; Glycoprotein; Signal.
KW	FT SIGNAL
FT	1 28
PROPEP	30
FT	POTENTIAL.

```

FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 162 AA; 18437 MW; 7EF7992391883446 CRC64;

Query Match 13.9%; Score 79.5; DB 1; Length 162;
Best Local Similarity 24.4%; Pred. No. 0.35;
Matches 29; Conservative 19; Mismatches 44; Indels 27; Gaps 6;

OY 5 IVDOLKYNVDLPDEF-----LPAPEDVETNCENSAFSCF-----OKAQLKSANTG 50
Db 53 VISDLKK-IEDILRSIHMDATLYTESDAHPNCKVTAKKCFLLERLVILQESRNSDSTV 111

OY 51 NNERII-NVSIKKLRKPPSTNAGRRQKRLTSPSCDSYKPKPKFLEFKSLIQKMI 108
Db 112 ENLIILANSLSIEYK--TESG-----CKECELEBKNEFLKFIHIVQWFI 159

RESULT 8
ALB2_XENLA STANDARD; PRT; 607 AA.
AC P14872:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE 74 KDA SERUM ALBUMIN PRECURSOR.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
SEQUENCE OF 3-607 FROM N.A.
RX MEDLINE=89313788; PubMed=2747653;
RA Moskatis J.E., Sargent T.D., Smith L.H. Jr., Pastori R.L.,
RA Schoenberg D.R.;
RT "Xenopus laevis serum albumin: sequence of the complementary
RT deoxyribonucleic acids encoding the 68- and 74-kilodalton peptides
RT and the regulation of albumin gene expression by thyroid hormone
RT during development.";
RL Mol. Endocrinol. 3:464-473(1989).
RN [2]
SEQUENCE OF 1-48 FROM N.A.
RX MEDLINE=88172470; PubMed=2451026;
RA Schorpp M., Doebbeling U., Wagner U., Ryffel G.U.;
RT "5'-flanking and 5'-proximal exon regions of the two xenopus albumin
RT genes. Deletion analysis of constitutive promoter function.";
RL J. Mol. Biol. 199:83-93(1988).
RN [3]
SEQUENCE OF 459-557 FROM N.A.
RX TISSUE=Liver;
RA MEDLINE=85126974; PubMed=3971963;
RA Wolffe A.P., Glover J.F., Martin S.C., Tenniswood M.P.R.,
RA Williams J.L., Tata J.R.;
RT "Deinduction of transcription of Xenopus 74-kda albumin genes and
RT destabilization of mRNA by estrogen in vivo and in hepatocyte
RT cultures.";
RL Eur. J. Biochem. 146:489-496(1985).
CC -1- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD
CC BINDING CAPACITY FOR WATER, CA++, NA+, K+, FATTY ACIDS, HORMONES,
CC BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
CC COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21442; AAA49637.1; -.
DR EMBL; M28276; AAA49642.1; -.
DR PIR; B41682; ABXL72.
DR HSSP; P02768; LUOR.
DR InterPro; IPR000264; -.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PRO0802; SERUMALBUMIN.
DR PROSITE; PS00212; ALBUMIN; 3.
KW Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; Signal;
KW Copper.
FT SIGNAL 1 18 POTENTIAL.
FT PROPEP 19 24 POTENTIAL.
FT CHAIN 25 607 74 KDA SERUM ALBUMIN.
FT REPEAT 29 211 1.
FT REPEAT 217 403 2.
FT REPEAT 409 601 3.
FT METAL 30 30 COPPER (BY SIMILARITY).
FT DISULFID 80 88 BY SIMILARITY.
FT DISULFID 101 117 BY SIMILARITY.
FT DISULFID 116 127 BY SIMILARITY.
FT DISULFID 147 192 BY SIMILARITY.
FT DISULFID 191 200 BY SIMILARITY.
FT DISULFID 223 269 BY SIMILARITY.
FT DISULFID 268 276 BY SIMILARITY.
FT DISULFID 288 302 BY SIMILARITY.
FT DISULFID 301 312 BY SIMILARITY.
FT DISULFID 339 384 BY SIMILARITY.
FT DISULFID 383 392 BY SIMILARITY.
FT DISULFID 415 461 BY SIMILARITY.
FT DISULFID 460 471 BY SIMILARITY.
FT DISULFID 484 500 BY SIMILARITY.
FT DISULFID 499 510 BY SIMILARITY.
FT DISULFID 537 582 BY SIMILARITY.
FT DISULFID 581 590 BY SIMILARITY.
FT CONFLICT 503 503 S -> L (IN REF. 3).
FT CONFLICT 531 531 H -> D (IN REF. 3).
SQ SEQUENCE 607 AA; 70382 MW; 592BA4177A36B66B CRC64;

Query Match 13.7%; Score 78.5; DB 1; Length 607;
Best Local Similarity 24.6%; Pred. No. 1.9;
Matches 35; Conservative 18; Mismatches 48; Indels 41; Gaps 5;

OY 4 DIVDOLKYNVDLPDEF-----LPAPEDVETNCENSAFSCF-----OKAQLKSAN 48
Db 72 EINDPAKSCINDKTPCEKPVGTLFDFDKLCADPAVGNYEWSKECCAKQDPERAQCFAH 131

OY 49 TGNNERII-----NVSIKKLRK-----PPSTNAGRRQKRLTSPS 84
Db 132 RDHEITSIKPEETCKLLKEHPDILLSAFIHEEARHNPDLPPAVLALTQYHKLAHC 191

OY 85 CDSYEKKPKPKFLEFKSLQK 106
Db 192 CEEEDKE--KCFSEKMKQLMKQ 211

RESULT 9
IL15_MOUSE STANDARD; PRT; 162 AA.
ID IL15_MOUSE
AC P48346;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=WC/REJ X C57BL/6J; TISSUE=Bone marrow;
RX MEDLINE=95278940; PubMed=7759105;
RA Anderson D.M., Johnson L., Glaccum M.B., Copeland N.G.,
RA Gilbert D.J., Jenkins N.A., Valentine V., Kirstein M.N.,
RA Shapiro D.N., Morris S., Grabstein K., Cosman D.,
RT "Chromosomal assignment and genomic structure of IL15."
RL Genomics 25:701-706(1995).
CC
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U14332; AAA75377.1; -.
DR MGD; MGI:103014; IL15.
DR Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48 POTENTIAL.
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 108 108 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SEQUENCE 162 AA; 18593 MW; 68C971498CEBF296 CRC64;
SQ
Query Match 12.7%; Score 72.5; DB 1; Length 162;
Best Local Similarity 20.0%; Pred. No. 1-7;
Matches 22; Conservative 29; Mismatches 54; Indels 5; Gaps 3;
QY 3 IDI---VDQLKNVNDL-VPEFLPAPEDVETNCWSAFSCFQKAQLKSANTGNRIINV 58
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 51 IDVRYDLKIESLQSIHIDTLYTDSDFHPSCVKVAMNCF-LLEQLVILHEYSNMTLNE 109
QY 59 SIKKLRKPPSTNAGRRXKRLTCPSDSEYKPKPFERFLKSLQKMI 108
:: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 110 TVRNVLVLANSTLSSNKNVAESGCKELEEKTFTEFLQSFIQVMFI 159
RESULT 10
Y069_NPVAC
ID Y069_NPVAC STANDARD; PRT; 262 AA.
AC F41469;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE HYPOTHETICAL 30.4 KDA PROTEIN IN LEF3-1AP2 INTERGENIC REGION.
OS Autographa californica nuclear polyhedrosis virus (AcMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae.
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=46015;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C6;
RX MEDLINE=94303173; PubMed=8030224;
RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT "The complete DNA sequence of Autographa californica nuclear
RL polyhedrosis virus."
RL Virology 202:586-605(1994).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; M25552; AAA70332.1; -;
 CC DR EMBL; X87611; CAA60929.1; -;
 CC DR EMBL; Z49507; CAA89529.1; -;
 CC DR PIR; A32108; A32108;
 CC DR SGD; S0003767; SU12;
 CC DR InterPro; IPR003029; -;
 CC DR Pfam; PF00575; S1; 1;
 CC KW Initiation factor; Protein biosynthesis; RNA-binding;
 KW Phosphorylation.
 FT MOD_RES 52 52 PHOSPHORYLATION (BY GCN2).
 SQ SEQUENCE 304 AA; 34717 MW; AF4FIC80303A4E98 CRC64;

Query Match 12.6%; Score 72; DB 1; Length 304;
 Best Local Similarity 33.3%; Pred. No. 3-9;
 Matches 20; Conservative 14; Mismatches 22; Indels 4; Gaps 4;
 QY 4 DIVDOLKNYND-LVPEFLPAPEDVETNC-EWSAFSCFQAKLKSANTGNRINVSIIK 61
 Db 164 DVLDELKNYSIKRLTPQAVKIRADVEVSCFSYEGIDAIDKA-LKSAEDMSTEQ-MQVKVK 221
 ::::: | | : | | | : | | | : | : | |

RESULT 12
 ID YIM7_YEAST STANDARD; PRT; 195 AA.
 AC P32459;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE UREIDOGLYCOLATE HYDROLASE (EC 3.5.3.19).
 GN DAL3 OR YIRO32C.
 OS Saccharomycetes cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 RN NCBI_TaxID=4932;
 RX [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE-92133160; PubMed-1776360;
 RA Yoo H.S., Cooper T.G.;
 RT "The ureidoglycolate hydrolase (DAL3) gene in Saccharomycetes cerevisiae.";
 RL Yeast 7:693-698(1991).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
 RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
 RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
 RA Walsh S.V., Whitehead S.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: UTILIZATION OF PURINES AS SECONDARY NITROGEN SOURCES,
 CC WHEN PRIMARY SOURCES ARE LIMITING.
 CC -1- CATALYTIC ACTIVITY: (-)-UREIDOGLYCOLATE + H(2)O = GLYOXYLATE +
 CC 2 NH(3) + CO(2).
 CC -1- PATHWAY: THIRD STEP IN THE DEGRADATION OF ALLANTOIN (PURINE
 CC CATABOLISM).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M64778; AAA73025.1; -;
 CC DR EMBL; Z38061; CAA86192.1; -;
 CC PIR; S42022; S42022.

DR PIR; S48494; S48494.
 DR SGD; S0001471; DAL3.
 KW Hydrolase; Purine metabolism; Prenylation; Lipoprotein.
 FT LIPID 192 FARNESYL (BY SIMILARITY).
 SQ SEQUENCE 195 AA; 21727 MW; 6230AEE69585206B CRC64;

Query Match 12.5%; Score 71.5; DB 1; Length 195;
 Best Local Similarity 29.2%; Pred. No. 2-7;
 Matches 19; Conservative 9; Mismatches 24; Indels 13; Gaps 1;

QY 6 VDOLKNYNDLVPEFLPAPEDVETNC-EWSAFSCFQAKLKSANTGNRINVSIIKLR 65
 Db 48 VSQVENKSTSKVP-----NWNLFRCFPQPHLNRVFTQGSNOAISHSIKVLK 94
 ::::: | | : | | | : | | | : | : | |

QY 66 KPPST 70
 Db 95 HPCST 99

RESULT 13
 ID YIM7_YEAST STANDARD; PRT; 206 AA.
 AC P40470;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HYPOTHETICAL 23.8 KDA PROTEIN IN MET18-SPH1 INTERGENIC REGION.
 GN YIL127C.
 OS Saccharomycetes cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 RN NCBI_TaxID=4932;
 RX [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
 RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
 RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
 RA Walsh S.V., Whitehead S.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; Z46833; CAA86865.1; -;
 CC DR SGD; S0001389; YIL127C.
 KW Hypothetical protein.
 SQ SEQUENCE 206 AA; 23845 MW; 4E51B11216F7ECB8 CRC64;

Query Match 12.1%; Score 69; DB 1; Length 206;
 Best Local Similarity 26.0%; Pred. No. 5;
 Matches 19; Conservative 14; Mismatches 40; Indels 0; Gaps 0;

QY 8 QLKKNYNDLVPEFLPAPEDVETNC-EWSAFSCFQAKLKSANTGNRINVSIIKLRKP 67
 Db 12 QATSVVGNLLSLLPVGVPKIRANNGKTSVYNGSKAQLIDRLNLRVQLQNRVHKIKKK 71
 QY 68 PSTNAGRQKHRL 80
 Db 72 KLVKKKKVKKHKL 84
 ::::: | | : | | | : | | | : | : | |

RESULT 14
 IDH_AZOVI

```

ID IDH_AZOV1 STANDARD; PRT: 741 AA.
AC .P16100;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE ISOCITRATE DEHYDROGENASE [NADP] (EC 1.1.1.42) (OXALOSUCCINATE
DE DECARBOXYLASE) (IDH).
GN ICD.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RA Sahara T., Suzuki M., Tsuruha J.I., Takada Y., Abe K., Fukunaga N.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP PRELIMINARY SEQUENCE OF 228-250 AND 254-259.
RX MEDLINE=74086945; PubMed=4149369;
RA Edwards D.J., Heinrikson R.L., Chung A.F.;
RT "Triphosphopyridine nucleotide specific isocitrate dehydrogenase from
RT Azotobacter vinelandii. Alkylation of a specific methionine residue
RT and amino acid sequence of the peptide containing this residue.";
RL Biochemistry 13:677-683(1974).
CC -!- CATALYTIC ACTIVITY: ISOCITRATE + NADP(+) = 2-OXOGLUTARATE +
CC CO(2) + NADPH.
CC -!- ENZYME REGULATION: INHIBITION OF THIS ENZYME BY PHOSPHORYLATION
CC REGULATES THE BRANCH POINT BETWEEN THE KREBS CYCLE AND THE
CC GLYOXYLATE BYPASS, WHICH IS AN ALTERNATE ROUTE THAT ACCUMULATES
CC CARBON FOR BIOSYNTHESIS WHEN ACETATE IS THE SOLE CARBON SOURCE
CC FOR GROWTH. THE PHOSPHORYLATION STATE OF THIS ENZYME IS CONTROLLED
CC BY ISOCITRATE DEHYDROGENASE KINASE/PHOSPHATASE (ACEK).
CC -!- SUBUNIT: MONOMER.
CC -!- SIMILARITY: BELONGS TO THE MONOMERIC-TYPE FAMILY OF IDH.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D73443; BRA11169.1; -
DR PIR; A10759;
KW Oxidoreductase; NADP; Phosphorylation; Glyoxylate bypass;
KW Tricarboxylic acid cycle.
SQ SEQUENCE 741 AA; 80389 MW; 29FF35278E5AED8B CRC64;

Query Match 12.1%; Score 69; DB 1; Length 741;
Best Local Similarity 26.1%; Pred. No. 20;
Matches 24; Conservative 18; Mismatches 28; Indels 22; Gaps 4;

QY 6 VDOLKNYNDL-----VPEFLPAPEDVETNCWESAFSCFQAOLKSAN-----TGNNEI 55
Db 90 VPOLAAALKELOOQGYKLPDY---PEEPKTDPEKDVAKYDKIGSAVNPVLREGNSDRR 146
QY 56 INVSIKILKRPPSTNAGRRQKRLTCTPCSDS 87
Db 147 APLSVKNYARKHP-----HKMGANSADS 169

RESULT 15
SDC2_HUMAN
ID SDC2_HUMAN STANDARD; PRT: 201 AA.
AC S34741.
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SYNDSCAN-2 PRECURSOR (FIBROGLYCAN) (HEPARAN SULFATE PROTEOGLYCAN CORE
DE PROTEIN) (HSPG) (SYND2).
GN SDC2 OR HSPG1.

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP NCBI_TaxID=9606;
RC SEQUENCE FROM N.A.
RX MEDLINE=89214123; PubMed=2523388;
RA Marynen P., Zhang J., Cassiman J.J., den Berghe H., David G.;
RT "Partial primary structure of the 48- and 90-kilodalton core proteins
RT of cell surface-associated heparan sulfate proteoglycans of lung
RT fibroblasts. Prediction of an integral membrane domain and evidence
RT for multiple distinct core proteins at the cell surface of human lung
RT fibroblasts.";
RL J. Biol. Chem. 264:7017-7024(1989).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE SYNDSCAN FAMILY OF INTEGRAL MEMBRANE
CC PROTEOGLYCANS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J04621; AAA52701.1; ALT_INIT.
DR MIM; 142460;
DR InterPro; IPR001050;
DR Pfam; PF01034; Syndscan; 1.
DR PROSITE; PS00964; SYNDSCAN; 1.
KW Proteoglycan; Heparan sulfate; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 201 SYNDSCAN-2.
FT DOMAIN 19 144 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 145 169 POTENTIAL.
FT SITE 170 201 CYTOPLASMIC (POTENTIAL).
FT SITE 142 143 CLEAVAGE OF ECTODOMAIN (POTENTIAL).
FT CARBOHYD 41 41 O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).
FT CARBOHYD 55 55 O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).
FT CARBOHYD 57 57 O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).
SQ SEQUENCE 201 AA; 22174 MW; 7B7F175650841054 CRC64;

Query Match 11.9%; Score 68; DB 1; Length 201;
Best Local Similarity 24.5%; Pred. No. 6.1;
Matches 24; Conservative 19; Mismatches 41; Indels 14; Gaps 2;

QY 7 DQKKNYNDLVPEFLPAPEDVETNCWESAFSCFQAOLKSANTGNNE-----RIINV 58
Db 108 DKEKVLHSDSERKMDPAEDTINVYTEKHSDSLFKRTEVLAAGVIGVFLFAIFLILL 167
QY 59 SIKKLARKPPSTNAGRRQKRLTCTPCSDSYEKPPKEF 96
Db 168 VYPMRKDEGSYDLGERK-----PSSAAYQKAPTKEF 199

Search completed: May 23, 2001, 11:20:23
Job time: 521 sec

```

Result No.	Score	Query Match	Length	DB	ID	Description
1	98.5	17.3	114	4	Q9UBA3	Q9ubA3 homo sapien
2	98.5	17.3	136	4	O00440	O00440 homo sapien
3	88.5	15.5	182	6	Q9XSJ6	Q9xsJ6 ovis aries
4	82	14.4	567	11	Q05208	Q05208 mus musculus
5	80.5	14.1	3052	14	Q82933	Q82933 johnsongrass
6	75	13.1	566	5	Q9VVE8	Q9vve8 drosophila
7	74	13.0	2197	5	O96296	O96296 plasmodium
8	73.5	12.9	187	13	Q9W756	Q9w756 gallus gall
9	73.5	12.9	868	5	Q9NAH8	Q9nah8 caenorhabdi
10	72.5	12.7	1101	5	O62471	O62471 caenorhabdi
11	72	12.6	262	14	Q92434	Q92434 bombyx mori
12	70.5	12.3	848	4	O75141	O75141 homo sapien
13	69	12.1	1220	10	O42015	O42015 lycopersico
14	68.5	12.0	336	11	O62612	O62612 rattus norv
15	68.5	12.0	566	11	O62611	O62611 rattus norv
16	68.5	12.0	1257	10	O64516	O64516 arabidopsis
17	68.5	12.0	1366	10	Q9LWQ6	Q9lmq6 arabidopsis
18	68.5	12.0	1670	5	O23901	O23901 dictyosteli
19	68.5	12.0	2873	14	Q93069	Q93069 hepatitis q

DT 01-JUL-1997 (TReMBLrel. 04, Created)

```

DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)
DE INTERLEUKIN-15.
GN IL-15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RL Meazza R., Ferrini S.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y09908; CAA71044.1; -
SQ SEQUENCE 136 AA; 15259 MW; 494ED8DDC5D364CD CRC64;

Query Match 17.3%; Score 98.5; DB 4; Length 136;
Best Local Similarity 26.1%; Pred. No. 0.0046;
Matches 30; Conservative 22; Mismatches 48; Indels 15; Gaps 5;

QY 3 IDIVDLKKNYNDLVPFF-----LPAPEDVTNCEWGAFCF----QKAQLKSANTGNNE 53
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
25 VNVISDLKK-IEDLIQSMHDATLYTSDVHPSCVKVAMKCFLELQVLISLESGDASHD 83
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 54 RIINVSYKKLKRPPSNAGRRQKHRLTCPSCDSEYKPKPEFLERKSLQKMI 108
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
84 TVENLI--LANNSSNGNVTES---GCKBCELEEKNIKEFLQSFVHVQMF 133
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
Q9XSJ6 PRELIMINARY; PRT; 162 AA.
ID Q9XSJ6
AC Q9XSJ6;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
DE INTERLEUKIN-15.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RL Casey G.J., Chaplin P.J.;
RL "Isolation of interleukin-15 mRNA transcripts from T and B cells
RL circulating in effertent lymph.";
RL J. Interferon Cytokine Res. 0:0-0(1999).
DR EMBL; AF149700; AAD37425.1; -
SQ SEQUENCE 162 AA; 18510 MW; 87195650E45F1E5D CRC64;

Query Match 15.5%; Score 88.5; DB 6; Length 162;
Best Local Similarity 28.9%; Pred. No. 0.063;
Matches 24; Conservative 10; Mismatches 48; Indels 1; Gaps 1;

QY 26 DVETNCFWSAFSCFQKALKSANTGNNERIINVSIIKKLKRPPSTNAGRRQKHRLTCPS 85
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
78 DAHPNCKVTALQCFLELRVLILHESKNAAYEI-IENITMLADRNLSIENKTELGCCKE 136
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 86 DSYEKKPKPEFLERKSLQKMI 108
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
137 EELEKKSIIKEFLKSFVHVQMF 159
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
Q05208 PRELIMINARY; PRT; 567 AA.
ID Q05208
AC Q05208;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE ST2L PROTEIN PRECURSOR.

```



```

DR INTERPRO: IPR001592;
DR INTERPRO: IPR001730;
DR INTERPRO: IPR002540;
DR PFAM: PF00270; DEAD; 1.
DR PFAM: PF00680; RNA_dep_RNA_pol; 1.
DR PFAM: PF00767; Poty_coat; 1.
DR PFAM: PF00851; Peptidase_C6; 1.
DR PFAM: PF00863; Peptidase_C4; 1.
DR PFAM: PF01577; Poty_P1; 1.
DR PRINTS: PR00966; NIAPORPTASE.
KW Coat protein; Protease.
FT CHAIN 239 699 HELPER COMPONENT (HC).
FT CHAIN 700 1046 PROPEASE3.
FT CHAIN 1047 1098 6K1.
FT CHAIN 1099 1757 CYTOPLASMIC INCLUSION (CI).
FT CHAIN 1758 1803 6K2.
FT CHAIN 1804 1991 SMALL NUCLEAR INCLUSION/VIRUS ATTACHED
FT CHAIN 1992 2233 PROTEIN.
FT CHAIN 2234 2749 SMALL NUCLEAR INCLUSION/PROTEASE.
FT CHAIN 2750 3052 LARGE NUCLEAR INCLUSION/POLYMERASE.
FT CHAIN 3052 347249 MW; 08CD8831A73EBCA9 CRC64;
SQ SEQUENCE 3052 AA; 347249 MW; 08CD8831A73EBCA9 CRC64;

Query Match 14.1%; Score 80.5; DB 14; Length 3052;
Best Local Similarity 23.1%; Pred. No. 9.9;
Matches 24; Conservative 22; Mismatches 55; Indels 3; Gaps 3;

QY 5 IVDOLKNNVNDLVPEFLPAPEDVETNCWSAFSCFQKALKSANTGNNGNRIINVSIIKLK 64
Db 220 LVNLDQVEED-VKQICHYSFADARAFWKGFETENHTAQREADHTNHEPV-MSVERCG 277
QY 65 RKPFSTNAGRQKRLTQPC-DSYEKKPPKPEFLERFKSLQKM 107
Db 278 RRAAMLENAFHOGFKITCKHCFQIFDEHSDEVCERIINALQRI 321

RESULT 6
QYVYE8 PRELIMINARY; PRT; 566 AA.
AC QYVYE8
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE CGI5745 PROTEIN.
GN CGI5745.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Eukaryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos J.B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

```

```

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner K., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003492; AAF48250.1; -.
DR FLYBASE; FBgn0030469; CGI5745.
SQ SEQUENCE 566 AA; 60620 MW; 31FA41FA4AEF427F CRC64;

Query Match 13.1%; Score 75; DB 5; Length 566;
Best Local Similarity 24.7%; Pred. No. 6.3;
Matches 20; Conservative 19; Mismatches 36; Indels 6; Gaps 2;

QY 26 DVETNCWSAF--CFQKALKSANTGNNGNRIINVSIIKLKRPSTNAGRQKRLTQPC 83
Db 120 DIGYCEVASISLALRKAQLKAQFFGNQVG----GLARDSETSTTRITRTTNRSAYP 175
QY 84 SCDSEYKPKPKPEFLERFKSL 104
Db 176 SCKTERGKPVQQLIDQFOAMI 196

RESULT 7
QYVYE8 PRELIMINARY; PRT; 2197 AA.
AC QYVYE8
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE PFEMP1
GN PFBI055C.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99021743; PubMed=9804551;
RA Gardner M.J., Tettelin H., Carucci D.J., Cummings L.M., Aravind L.,
RA Koonin E.V., Shalloo S., Mason T., Yu K., Fujii C., Pederson J.,
RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M.,
RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Science 282:1126-1132(1998).
DR EMBL; AE001434; AAC71996.1; -.
SQ SEQUENCE 2197 AA; 249668 MW; D2A92DB54535C143 CRC64;

Query Match 13.0%; Score 74; DB 5; Length 2197;
Best Local Similarity 29.5%; Pred. No. 34;
Matches 28; Conservative 16; Mismatches 41; Indels 10; Gaps 4;

QY 9 LKXNYNDLVPEFLPAPEDVETNCWSAF--SCFQKALKSANTGNNGNRIINVSIIKLK 66
Db 1594 VKSFLETWPK-IAVVNDQDNVILKSKFGNSCGCSAISTN-GNEBDAIDCMIKLEKK 1651

```

```

Qy 67 -----PPSTNAGRQKRLTCTPCSDSYEKKPKPE 95
      | | | | | | | | | | | | | | | | | |
Db 1652 IDECKRRPGSGQTCNETLTHPLDVQDEPLEE 1686

RESULT 8
Q9W756 PRELIMINARY; PRT; 187 AA.
AC Q9W756;
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (INTERLEUKIN 15 PRECURSOR),
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
RX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCWL; TISSUE=LIVER;
RA Burnside J., Sofer L.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=SC; TISSUE=LIVER;
RA Choi K.D., Lillehoj H.S., Burnside J.;
RL "Gallus gallus mrna for IL-15 precursor."
RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF152927; AAD38392.1;
DR EMBL: AF139097; AAF61446.1;
SQ SEQUENCE 187 AA; 21964 MW; 470601EBF8837095 CRC64;

Query Match 12.9%; Score 73.5; DB 13; Length 187;
Best Local Similarity 25.0%; Pred. No. 2.8;
Matches 24; Conservative 11; Mismatches 34; Indels 27; Gaps

Qy 25 EDVETNCWSAFSCF---QKAQLKSANTGNRIINVTIKLKRKPPSTNAGRQKRLT 81
      ||| | | | | | | | | | | | | | | |
Db 98 EDIE--CQEPVMRCFFLEMLVILHECD-----IKKCRKHQDVRIWNGNARFA 144

Qy 82 -----CPSCDSYEKKPKPELRFKSLQK 106
      | | | | | | | | | | | | | | | | | |
Db 145 TYQLNSTTAKCKECEYEENKTFIQSFVKVIQR 180

RESULT 9
Q9NAH8 PRELIMINARY; PRT; 868 AA.
AC Q9NAH8;
DT 01-OCT-2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
DE Y47D3A.14 PROTEIN.
GN Y47D3A.14
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
RX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Matthews L.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: AL117202; CAB55073.1;
SQ SEQUENCE 868 AA; 100165 MW; 09A2A97066E373D7 CRC64;

```

DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 DE ACNPV ORF69.
 OS Bombyx mori nuclear polyhedrosis virus (BmNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10458;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T3;
 RX MEDLINE=97329351; PubMed=9185864;
 RA Kamita S.G., Maeda S.;
 RT "Sequencing of the putative DNA helicase-encoding gene of the Bombyx mori nuclear polyhedrosis virus and fine-mapping of a region involved in host range expansion."
 RT Gene 190:173-179(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=T3;
 RA Goni S., Majima K., Maeda S.;
 RT "Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus."
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: L33180; AAC63742.1; -
 DR INTERPRO: IPR002877;
 DR PFAM: PF01728; Ftsu; 1.
 DR PRINTS: PR001728; Ftsu; 1.
 SQ SEQUENCE 262 AA; 30384 MW; E3AF85E1B02DA48C CRC64;

 Query Match 12.6%; Score 72; DB 14; Length 262;
 Best Local Similarity 21.7%; Pred. No. 5.8; Mismatches 19; Indels 10; Gaps 2;
 Matches 18; Conservative 36;

 QY 30 NCWSAFSCFOKALQSANTGNRIINVSIIKKLRKPPSTNAGRRQKH-----RLTCPS 84
 Db 176 NCVLKVFDAEHTQIOMLN-----KFNHFKEKVLKPPSSRPANSEYLLCFNKLVRPY 230

 QY 85 CDSYEKKPKPKFLEFRKSLQKM 107
 Db 231 CNDYVNELEKQFKYRIQLKNL 253

 RESULT 12
 ID 075141 PRELIMINARY; PRT; 848 AA.
 AC 075141;
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DE KIAA0650 PROTEIN (FRAGMENT).
 GN KIAA0650.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=98403880; PubMed=9734811;
 RA Ishikawa K., Nagase Y., Suyama M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. X. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro."
 RL DNA Res. 5:169-176(1998).
 DR EMBL: AB014550; BAA31625.1; -
 FT NON_TER
 SQ SEQUENCE 848 AA; 95756 MW; 2192522F40ED9E7B CRC64;

 Query Match 12.3%; Score 70.5; DB 4; Length 848;
 Best Local Similarity 25.9%; Pred. No. 29;
 Matches 21; Conservative 17; Mismatches 38; Indels 5; Gaps 3;

QY 19 EFLPAPEDVETNC-EWSAFSCFOKALQSANTGNRIINVSIIKKLRKPPSTNAGRRQK 77
 Db 66 KFIQPGPGNKDLCTWREFSDFIRVOLIS---GPPAKLLIDWPELKESIPVIN-GRDQ 121

 QY 78 HRLTCPSCDSEYKKPKPEFLE 98
 Db 122 NPIIVOLCDOWNDNPAPVQHVK 142

 RESULT 13
 ID 024015 PRELIMINARY; PRT; 1220 AA.
 AC 024015;
 DT 01-JAN-1998 (Tremblrel. 05, Created)
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
 DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
 DE RESISTANCE COMPLEX PROTEIN I2C-1.
 GN I2C-1.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97290204; PubMed=9144960;
 RA Ori N., Eshed Y., Paran I., Presting G., Aviv D., Tanksley S., Zhanir D., Fluhr R.;
 RT "The I2C family from the wilt disease resistance locus I2 belongs to the nucleotide binding, leucine-rich repeat superfamily of plant resistance genes."
 RT Plant Cell 9:521-532(1997).
 DR EMBL: AF004878; AAB63274.1; -
 DR MENDEL: 24472; Lyces; 3172; 24472.
 DR INTERPRO: IPR000767; -
 DR INTERPRO: IPR001611; -
 DR PFAM: PF00560; LRR; 4.
 DR PFAM: PF00931; NB-ARC; 1.
 DR PRINTS: PR00364; DISEASERSIST.
 SQ SEQUENCE 1220 AA; 139386 MW; B5997280C8C5EB20 CRC64;

 Query Match 12.1%; Score 69; DB 10; Length 1220;
 Best Local Similarity 30.3%; Pred. No. 61;
 Matches 30; Conservative 8; Mismatches 23; Indels 38; Gaps 6;

 QY 13 VNDLVPEFLPA-----PEDVET-----NCEWSAFSCFOKALQSANTGN 51
 Db 948 VDDISPEFLPTARQLSIENCHNVTRFLIPTATESLHNRCEKLSMACGGAAQLTSLNIW- 1006

 QY 52 NERIINVSIIKKLRKP---PSTNAGRRQKHRLT-CPSCD 86
 Db 1007 -----GCKKLKLPPELLPSL-----KELRLTYCPCEIE 1033

 RESULT 14
 ID 062612 PRELIMINARY; PRT; 336 AA.
 AC 062612;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE FIT-1S PRECURSOR.
 GN FIT-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94178260; PubMed=8131748;

GenCore version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:11:33 ; Search time 108.07 Seconds
(without alignments)
69.292 Million cell updates/sec

Title: US-09-522-217-2_COPY_32_162

Perfect score: 695

Sequence: 1 QDRHMRQLDIVDQLKN.....LLQKMIHOLSRTHGSEDS 131

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	695	100.0	162	21	A human zalphall 1
2	695	100.0	119	21	Amino acid sequenc
3	394	56.7	146	21	A mouse zalphall 1
4	394	56.7	510	21	Amino acid sequenc
5	212	30.5	40	21	Antigenic peptide
6	176	25.3	32	21	Antigenic peptide
7	103.5	14.9	135	21	Human Interleukin-
8	99.5	14.3	114	16	Human Interleukin-
9	99.5	14.3	114	16	Mammalian interleu
10	99.5	14.3	114	16	Mammalian IL-15.
11	99.5	14.3	114	17	Simian mature epit

12	99.5	14.3	114	17	W07253	Generic mammalian
13	99.5	14.3	114	19	W39186	Simian epithelium
14	99.5	14.3	114	20	Y03757	Simian epithelium
15	99.5	14.3	114	21	Y52309	Mature simian epit
16	99.5	14.3	162	16	R83438	Human interleukin-
17	99.5	14.3	162	16	R66927	Human IL-15. Homo
18	99.5	14.3	162	17	W09098	Simian epithelium
19	99.5	14.3	162	17	W07255	Human epithelium-d
20	99.5	14.3	162	17	R98527	Human interleukin-
21	99.5	14.3	162	18	W37369	Wild-type interleu
22	99.5	14.3	162	18	W01658	Human interleukin-
23	99.5	14.3	162	19	W53878	Human interleukin-
24	99.5	14.3	162	19	W39185	Simian epithelium
25	99.5	14.3	162	20	Y03756	Simian epithelium
26	99.5	14.3	162	21	B18632	Amino acid sequenc
27	99.5	14.3	162	21	Y78595	Human interleukin-
28	99.5	14.3	162	21	Y52308	Simian epithelium-
29	99.5	14.3	162	22	B50870	Human IL-15. Homo
30	94.5	13.6	114	16	R83309	Simian interleukin
31	94.5	13.6	114	17	W09101	Human mature epith
32	94.5	13.6	114	19	W39188	Human epithelium d
33	94.5	13.6	114	20	Y03759	Human epithelium-d
34	94.5	13.6	114	21	Y52311	Mature human epith
35	94.5	13.6	122	17	R90842	Recombinant flag s
36	94.5	13.6	162	16	R83436	Simian interleukin
37	94.5	13.6	162	16	R66926	Simian IL-15. Cer
38	94.5	13.6	162	17	W09100	Human epithelium d
39	94.5	13.6	162	17	W07254	Simian epithelium-
40	94.5	13.6	162	17	R98526	Simian interleukin
41	94.5	13.6	162	17	R92798	Mammalian interleu
42	94.5	13.6	162	18	W37370	Mutant interleukin
43	94.5	13.6	162	19	W39187	Human epithelium d
44	94.5	13.6	162	20	Y03758	Human epithelium-d
45	94.5	13.6	162	21	Y78594	Simian interleukin

ALIGNMENTS

RESULT 1

ID	B18623	standard; Protein; 162 AA.
AC	B18623;	
DT	22-JAN-2001	(first entry)
DE	A human zalphall ligand polypeptide.	
XX	zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;	
KW	tumorigenesis; leukaemia; hematopoiesis; B cell tumour.	
XX	Homo sapiens.	
OS		
XX		
PN	W0200053761-A2.	
XX		
PD	14-SEP-2000.	
XX		
PF	09-MAR-2000; 2000WO-US06067.	
XX		
PR	09-MAR-1999; 99US-0264908.	
PR	11-MAR-1999; 99US-0265992.	
PR	01-JUL-1999; 99US-0142013.	
XX	(ZYMO) ZYMOGENETICS INC.	
XX		
PI	Novak JE, Freshnell SR, Sprecher CA, Foster DC, Gross JA;	
PI	Johnston JV, Nelson AJ, Dillon SR, Hammond AK;	
XX		
XX	WPI: 2000-565600/52.	
DR	N-PSDB: A75552.	
XX		
PT	New human cytokine, designated zalphall ligand, useful for stimulating	

PT the proliferation and/or development of haematopoietic cells in vitro
 XX and in vivo, and for treating tumourigenesis -
 PS Disclosure; Page 205-206; 256pp; English.
 XX The present sequence represents a human zalphall ligand polypeptide,
 CC which is a cytokine. The zalphall ligand is useful for stimulating the
 CC proliferation and development of haematopoietic cells in vitro and in
 CC vivo. Zalphall ligand polynucleotides can be used as primers or probes
 CC for cloning the zalphall gene. The zalphall ligand is useful for
 CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
 CC used for treating leukaemias and lymphomas. Antagonists against zalphall
 CC ligand are useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 SQ Sequence 162 AA;

Query Match 100.0%; Score 695; DB 21; Length 162;
 Best Local Similarity 100.0%; Pred. No. 8.5e-73;
 Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDRHMRMROLIDIVDOLKNVNDLVPELPAPEDVETNCESAFSCFOKAOLKSANTGN 60
 DB 32 qdrhmrmrqlidivdqiknyvndlvpefpapedvetncesafscfkaqlksantgn 91
 QY 61 NERIINVSIIKKLRKPPSTNAGRQKRLTCPCSDSYEKKKPKFLEFKSLQKMIHOH 120
 DB 92 neriinvsikkkrppstnagrrqkrltccpscdsyekkkpkflefekslqkmiqh 151
 QY 121 LSSRTHGSEDS 131
 DB 152 lssrthgseds 162

RESULT 2
 B18627
 ID B18627 standard; Protein; 519 AA.
 XX
 AC B18627;
 XX
 DT 22-JAN-2001 (first entry)
 XX
 DE Amino acid sequence of MBP-human zalphall ligand fusion in pTAP126.
 XX
 KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO200053761-A2;
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06067.
 XX
 PR 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 XX
 XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 XX

DR WPI; 2000-565600/52.
 XX N-PSDB; A75599.
 PT New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 XX and in vivo, and for treating tumourigenesis -
 PS Example 31; Page 233-235; 256pp; English.
 XX The present sequence represents a MBP-human zalphall ligand fusion in
 CC the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
 CC is useful for stimulating the proliferation and development of
 CC haematopoietic cells in vitro and in vivo. Zalphall ligand
 CC polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 SQ Sequence 519 AA;

Query Match 100.0%; Score 695; DB 21; Length 519;
 Best Local Similarity 100.0%; Pred. No. 3.9e-72;
 Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDRHMRMROLIDIVDOLKNVNDLVPELPAPEDVETNCESAFSCFOKAOLKSANTGN 60
 DB 389 qdrhmrmrqlidivdqiknyvndlvpefpapedvetncesafscfkaqlksantgn 448
 QY 61 NERIINVSIIKKLRKPPSTNAGRQKRLTCPCSDSYEKKKPKFLEFKSLQKMIHOH 120
 DB 449 neriinvsikkkrppstnagrrqkrltccpscdsyekkkpkflefekslqkmiqh 508
 QY 121 LSSRTHGSEDS 131
 DB 509 lssrthgseds 519

RESULT 3
 B18624
 ID B18624 standard; Protein; 146 AA.
 XX
 AC B18624;
 XX
 DT 22-JAN-2001 (first entry)
 XX
 DE A mouse zalphall ligand polypeptide.
 XX
 KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX
 OS Mus musculus.
 XX
 PN WO200053761-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06067.
 XX
 PR 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA


```

XX PF 09-MAR-2000; 2000WO-US06067.
XX XX
XX PR 09-MAR-1999; 99US-0264908.
XX PR 11-MAR-1999; 99US-0265992.
XX PR 01-JUL-1999; 99US-0142013.
XX XX
XX PA (ZYMO ) ZYMOGENETICS INC.
XX XX
XX PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
XX PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX DR WPI; 2000-565600/52.
XX XX
XX PT New human cytokine, designated zalpall ligand, useful for stimulating
XX PT the proliferation and/or development of haematopoietic cells in vitro
XX PT and in vivo, and for treating tumourigenesis -
XX PS Example 34; Page 227; 256pp; English.
XX XX
XX CC The present sequence was used to raise antibodies, and is derived from
XX CC a human zalpall ligand polypeptide. Zalpall ligand is a cytokine.
XX CC The zalpall ligand is useful for stimulating the proliferation and
XX CC development of haematopoietic cells in vitro and in vivo. Zalpall
XX CC ligand polynucleotides can be used as primers or probes for cloning the
XX CC zalpall gene. The zalpall ligand is useful for treating tumourigenesis.
XX CC A zalpall ligand-saporin fusion toxin may be used for treating
XX CC leukaemias and lymphomas. Antagonists against zalpall ligand are
XX CC useful as research reagents for characterizing ligand-receptor
XX CC interaction. Antagonists are also useful for inhibiting expansion,
XX CC proliferation, activation and differentiation of cells involved in
XX CC regulating hematopoiesis. The zalpall ligand may also be used to
XX CC stimulate an immune response against B cell tumour, a virus, a parasite
XX CC or a bacterium. The zalpall polypeptides, polynucleotides, antagonists,
XX CC agonists and antibodies are also useful for the detection, diagnosis,
XX CC prevention, and treatment of diseases associated with a zalpall ligand
XX CC genetic defect.
XX SQ Sequence 40 AA;

Query Match 30.5%; Score 212; DB 21; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.4e-17;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QDRHMRQLIDIVDLKKNYNDLVPEFLPAPEDVETNC 40
Db 1 qdrhmrmqlidivdlkknynyndlvpeflpapedvetnc 40

RESULT 6
B18626
ID B18626 standard; Peptide; 32 AA.
XX AC B18626;
XX XX
XX DT 22-JAN-2001 (first entry)
XX DE Antigeninc peptide derived from a human zalpall ligand polypeptide.
XX KW zalpall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
XX KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
XX XX
XX OS Homo sapiens.
XX PN WO200053761-A2.
XX XX
XX PD 14-SEP-2000.
XX XX
XX PF 09-MAR-2000; 2000WO-US06067.
XX PR 09-MAR-1999; 99US-0264908.
XX PR 11-MAR-1999; 99US-0265992.

```

```

PR 01-JUL-1999; 99US-0142013.
XX (ZYMO ) ZYMOGENETICS INC.
XX XX
XX PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
XX PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX XX
XX DR WPI; 2000-565600/52.
XX XX
XX PT New human cytokine, designated zalpall ligand, useful for stimulating
XX PT the proliferation and/or development of haematopoietic cells in vitro
XX PT and in vivo, and for treating tumourigenesis -
XX PS Example 34; Page 227; 256pp; English.
XX XX
XX CC The present sequence was used to raise antibodies, and is derived from
XX CC a human zalpall ligand polypeptide. Zalpall ligand is a cytokine.
XX CC The zalpall ligand is useful for stimulating the proliferation and
XX CC development of haematopoietic cells in vitro and in vivo. Zalpall
XX CC ligand polynucleotides can be used as primers or probes for cloning the
XX CC zalpall gene. The zalpall ligand is useful for treating tumourigenesis.
XX CC A zalpall ligand-saporin fusion toxin may be used for treating
XX CC leukaemias and lymphomas. Antagonists against zalpall ligand are
XX CC useful as research reagents for characterizing ligand-receptor
XX CC interaction. Antagonists are also useful for inhibiting expansion,
XX CC proliferation, activation and differentiation of cells involved in
XX CC regulating hematopoiesis. The zalpall ligand may also be used to
XX CC stimulate an immune response against B cell tumour, a virus, a parasite
XX CC or a bacterium. The zalpall polypeptides, polynucleotides, antagonists,
XX CC agonists and antibodies are also useful for the detection, diagnosis,
XX CC prevention, and treatment of diseases associated with a zalpall ligand
XX CC genetic defect.
XX SQ Sequence 32 AA;

Query Match 25.3%; Score 176; DB 21; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.6e-13;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 91 CPSCDSYEKKPKPEFLERFKSLQKMIHQHLS 122
Db 1 cpscdsyekppkeflerfkslqkmiqhls 32

RESULT 7
Y54825
ID Y54825 standard; Protein; 135 AA.
XX AC Y54825;
XX XX
XX DT 04-FEB-2000 (first entry)
XX DE Human Interleukin-15 protein sequence.
XX KW Interleukin-15; IL-15; antisense inhibitor; untranslated region; 5'UTR;
XX KW 3'UTR; human; infection; inflammation; tumour; therapy; diagnosis.
XX XX
XX OS Homo sapiens.
XX PN US5985663-A.
XX XX
XX PD 16-NOV-1999.
XX XX
XX PF 25-NOV-1998; 98US-0200141.
XX XX
XX PR 25-NOV-1998; 98US-0200141.
XX XX
XX PA (ISIS-) ISIS PHARM INC.
XX PI Bennett CF, Cowser LM;
XX XX
XX DR WPI; 2000-022283/02.

```

DR N-PSDB; Z37358.
XX Antisense compound useful for inhibiting human interleukin-15
PT expression useful for treating diseases associated with interleukin-15
PT expression -
XX Example 13; Column 43-44; 31pp; English.
XX This sequence is the human interleukin-15. The invention relates to
CC antisense compounds that are targeted to a 5' or 3' untranslated region
CC (5'UTR or 3'UTR) of a nucleic acid molecule encoding human interleukin-15
CC (IL-15), and inhibit the expression of human IL-15. The antisense
CC inhibitors are useful for inhibiting expression of IL-15 in human
CC cells or tissues in vitro, for treating humans or other animals suspected
CC of having or being prone to a disease associated with IL-15 expression,
CC e.g. infections, inflammation or tumours. The inhibitors can also be used
CC for research or diagnostic purposes. Using antisense compounds
CC specifically and effectively inhibits IL-15 function.
XX Sequence 135 AA;
SQ

Query Match 14.9%; Score 103.5; DB 21; Length 135;
Best Local Similarity 26.7%; Pred. No. 0.00026;
Matches 31; Conservative 22; Mismatches 48; Indels 15; Gaps 5;
QY 12 IDIVDLKKNVNDLPVEF-----LPAPEDVETNCESAFSCF-----OKAQLKSANTGNNE 62
Db 24 vnvisdllk-iedliqsmhidatlytesdvhpckvtamkcfillelqvlsesgdasihd 82
QY 63 RIINVSIKKLRKPPSTNAGRQKHRLTCPCSDSYEKKPKPEFLERKSLLOKMIH 118
Db 83 tvenlii--lannslsngnvtcs---gckeeceeleknikeflqsfvhiqvmfin 133

RESULT 8
R83310
ID R83310 standard; Protein; 114 AA.
AC
AC R83310;
DT 02-FEB-1996 (first entry)
DE Human interleukin-15 mature polypeptide.
KW Interleukin-15; epithelium-derived T-cell factor; T lymphocyte.
XX Homo sapiens.
XX
XX W09527722-A.
XX 19-OCT-1995.
XX 06-APR-1994; 94WO-US03793.
XX 06-APR-1994; 94WO-US03793.
XX (IMMV) IMMUNEX CORP.
XX Anderson DM, Eisenman JR, Fung V, Grabstein KH;
PI Rauch C;
XX WPI; 1995-373556/48.
XX N-PSDB; T00527.
XX Isolated DNA encoding polypeptide with mammalian IL-15 activity - which
PT stimulates proliferation and differentiation of T cells, used for
PT treating carcinoma(s), melanomas, etc. and viral infections
XX
XX Claim 23; Page 30; 48pp; English.
XX A simian species of IL-15 (sIL-15) was purified and its AA
CC sequence and cDNA sequence analysed (see R83309, R83436,

CC T00524, T00525). Both the simian and the human ORFs encode
CC a precursor polypeptide (R83436, R83438). The precursor
CC polypeptides each comprise a 48-AA leader sequence and a sequence
CC encoding mature simian or human IL-15 polypeptides. The active
CC simian and human IL-15 polypeptides are disclosed in R83309 &
CC R83310 respectively. The invention also comprises other mammalian
CC IL-15, including human IL-15, that hybridise to probes defined by
CC R83438. A plasmid contg. a recombinant clone of human IL-15
CC cDNA was deposited with ATCC on Feb. 19 1993 under ATCC 69245.
CC The deposit was named 141-hETF. R83435 is a mammalian mature
CC IL-15 polypeptide. It is a generic sequence which encompasses both
CC R83309 (simian) and R83310 (human) IL-15 mature polypeptides.
XX
SQ Sequence 114 AA;
Query Match 14.3%; Score 99.5; DB 16; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00062;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;
QY 12 IDIVDLKKNVNDLPVEF-----LPAPEDVETNCESAFSCF-----OKAQLKSANTGNNE 62
Db 3 vnvisdllk-iedliqsmhidatlytesdvhpckvtamkcfillelqvlsesgdasihd 61
QY 63 RIINVSIKKLRKPPSTNAGRQKHRLTCPCSDSYEKKPKPEFLERKSLLOKMIH 118
Db 62 tvenlii--lannslsngnvtcs---gckeeceeleknikeflqsfvhiqvmfin 112

RESULT 9
R83435
ID R83435 standard; Protein; 114 AA.
AC
AC R83435;
XX
XX 02-FEB-1996 (first entry)
XX Mammalian interleukin-15 mature polypeptide.
DE Interleukin-15; epithelium-derived T-cell factor; T lymphocyte.
XX
XX Mammalian.
XX
XX Key Location/Qualifiers
XX Misc-difference 52 /label= L,H
XX Misc-difference 57 /label= A,T
XX Misc-difference 58 /label= S,D
XX Misc-difference 73 /label= S,I
XX Misc-difference 80 /label= V,I
XX W09527722-A.
XX 19-OCT-1995.
XX 06-APR-1994; 94WO-US03793.
XX 06-APR-1994; 94WO-US03793.
XX (IMMV) IMMUNEX CORP.
XX Anderson DM, Eisenman JR, Fung V, Grabstein KH;
PI Rauch C;
XX WPI; 1995-373556/48.
XX Isolated DNA encoding polypeptide with mammalian IL-15 activity - which
PT stimulates proliferation and differentiation of T cells, used for
PT treating carcinoma(s), melanomas, etc. and viral infections
XX

XX	WPI; 1995-082473/11.
DR	
XX	
PT	New purified interleukin-15.- which induces T cell proliferation
PT	and differentiation, used for the treatment of tumours and viral
PT	infection
XX	
XX	Claim 1; Page 33; 47pp; English.
PS	
XX	
CC	Simian and human IL-15 cDNAs (Q84583-84) can be used to obtain
CC	cDNAs encoding other mammalian homologs of IL-15. A general
CC	sequence of mammalian IL-15 is claimed.
XX	
XX	Sequence 114 AA;
SQ	

```

Query Match      14.3%; Score 99.5; DB 16; Length 114;
Best Local Similarity 25.2%; Pred. No. 0.0062;
Matches 30; Conservative 23; Mismatches 45; Indels 21; Gaps 5;

Qy 12 IDIVDQLKNVNDLVPEF-----LPAPEDVETNCSEAFSCF-QKAOLKSANTGN--- 61
    : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 vnvisdikk-iedliqsmhidatlytstdvhpsckvtamckfllelqvixesgdxihd 61

Qy 62 --ERIIIVNSIKKLKRPSTNAGRRQRHRLTCSDSYEKKKPKPEFVERKSLLOKMIH 118
    : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 tvenliilannxissngnxxtesg-----ckeeceleeeknfelqsfvhwgmfin 112
    : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 11
W09099
ID W09099 standard; Protein: 114 AA.

AC	W090999;
AD	
AE	
AF	11-MAR-1997 (first entry)
AG	
AH	
AI	Simian mature epithelium derived T cell factor.
AJ	
AK	
AL	SETF; African green monkey; CV1/EBNA cell; T-cell; B-cell;
AM	lymphocyte; proliferation; differentiation; gastrointestinal;
AN	KW infection; human immunodeficiency virus.
AO	
AP	
AQ	Cercopithecus aethiops.
AR	
AS	
AT	
AV	Key
AW	Location/Qualifiers
AX	1..114
AY	/label= mature_SETF
AZ	
BA	US5574138-A.
BB	
BC	12-NOV-1996.
BD	
BE	
BF	08-MAR-1993; 93US-0031399.
BG	
BH	22-FEB-1995; 95US-0393305.
BI	08-MAR-1993; 93US-0031399.
BJ	22-APR-1994; 94US-0233606.
BK	
BL	(IMMV) IMMUNEX CORP.
BM	
BN	Anderson DM, Eisenman JR, Fung V, Grabstein KH;
BO	Rauch C;
BP	
BQ	
BR	WPI; 1996-517923/51.
BS	N-PSDB; T49455.
BT	
BU	
BV	New epithelium derived T cell factor - induces proliferation of T
BW	and B cells, stimulates destruction of tumour and virus-infected
BX	T cells and protects against toxicity, partic. for treating intestinal
BY	disease and HIV infection
BZ	
CA	Claim 1; Fig 1; 35pp; English.
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	
CC	
CD	
CE	
CF	
CG	
CH	
CI	
CJ	
CK	
CL	
CM	
CN	
CO	
CP	
CQ	
CR	
CS	
CT	
CU	
CV	
CW	
CX	
CY	
CA	
CB	</

Claim 1; Column 29; 22pp; English.
 This is the amino acid sequence of a generic mammalian epithelium-derived T cell factor (ETF). The pref. ETF are the simian (W07254) or human (W07255) proteins. This sequence represents the generic sequence of the mature protein. Both the human and simian proteins contain a 48 amino acid leader sequence. ETF is a protein of 15-17 kd which is expressed by epithelial cells and stimulates proliferation and/or differentiation of precursor and/or mature T cells. The protein is therefore useful for promoting long term in vivo culture of T-lymphocytes and T-cell lines.
 Sequence 114 AA;
 SQ

Query Match	14.3%	Score 99.5	DB 17	Length 114
Best Local Similarity	25.2%	Pred. No. 0.00062		
Matches	30	Conservative 23	Mismatches 45	Indels 21
Gaps	5			
Qy	12	IDIVDQLKNYNDLVPEF----	IPAPEDVTNEWSAFSCF-OQAOLKSANTGNN----	61
	:::			
Db	3	vnvisdlkk-iedliqsmhidatlytsdsvhpscvktamkcflllelqvlvixesgdxxihd	61	
Qy	62	--ERIIINVISKKLRKPPSTNAGRRQRHLTCPCSDSYEKKPPPEFLERFKSLLOKMIH	118	
Db	62	tvenliilannxlssngnxtesg-----	ckeceeleeknikelfqsfvhiqvgmin	112

RESULT	13
W39186	
ID	W39186 standard; Protein; 114 AA.
XX	
XX	
AC	W39186;
XX	
XX	
DT	08-MAY-1998 (first entry)
XX	
XX	
DE	Simian epithelium derived T-cell factor mature protein.
XX	
XX	
KW	Epithelium derived T-cell factor; ETF; simian; gastrointestinal disease;
KW	B-cell stimulation; epithelium; villi; CD4+; T-cell stimulation;
KW	treatment; prevention.
XX	
XX	
OS	Simian.

Key	Location/Qualifiers
1..114	
/label= ETF	
US5707616-A.	
13-JAN-1998.	
04-OCT-1996;	96US-0726817.
22-FEB-1995;	95US-0393305.
08-MAR-1993;	93US-0031399.
22-APR-1994;	94US-0233606.
(IMMV) IMMUNEX CORP.	
Anderson DM, Eisenman JR, Fung V, Grabstein KH; Rauch C;	
WPI; 1998-100295/09.	
N-PSDB; V02873.	
Treatment or prevention of gastrointestinal diseases - by administering epithelium-derived T-cell factor polypeptide	

This sequence represents a simian epithelium-derived T-cell factor (ETF) mature protein which is used in a method for treating or preventing

CC gastrointestinal disease. These polypeptides have participated with application
CC in the treatment of gastrointestinal disorders associated with disruption
CC of the gastrointestinal epithelium or villi such as chemotherapy- and
CC radiation-therapy induced enteritis (gut toxicity), mucositis, peptic
CC ulcer disease, gastroenteritis and colitis, villus atrophic disorders,
CC malignancy and inflammatory bowel disease. ETF polypeptides may also be
CC useful in the treatment of human immunodeficiency virus (HIV) and
CC HIV-associated disease due to their ability to stimulate CD4+ and CD8+
CC cells. Biologically active ETF may be used to treat a variety of other
CC diseases or conditions where T-cell or B cell stimulation is desired.

XX SQ Sequence 114 AA;

Query Match 14.3%; Score 99.5; DB 19; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00062;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSEAFSCF---QKAQLKSANTGNNE 62
Db 3 vnvvisdlkk-iedliqsmhidatlytesdvhpckvtamkcfllleqlvislesgasihd 61
QY 63 RIINVSIIKKLRKPPSTNAGRQKRLTCPCDSYKPKPEFLERFKSLLOKMIH 118
Db 62 tvenlii--lannslsngnvtes---gckeceeleeknikelfqsfvhiqvmfin 112

RESULT 14
Y03757
ID Y03757 standard; Protein; 114 AA.
XX AC Y03757;
XX DT 10-JUN-1999 (first entry)
XX DE Simian epithelium-derived T-cell factor (ETF) mature protein sequence.
XX KW Epithelium-derived T-cell factor; ETF; immunoassay; enteritis; ulcer;
KW T cell proliferation; gastrointestinal disease; mucositis; colitis;
KW gastroenteritis; villus atrophy; malignancy; inflammatory bowel disease;
KW human immune deficiency virus; tumour; simian.
XX OS Mammalia.
XX PN US5892001-A.
XX PD 06-APR-1999.
XX PF 04-OCT-1996; 96US-0725969.
XX PR 22-FEB-1995; 95US-0393305.
XX PR 08-MAR-1993; 93US-0031399.
XX PR 22-APR-1994; 94US-0233606.
XX PR 04-OCT-1996; 96US-0725969.
XX PA (IMMV) IMMUNEX CORP.
XX PI Anderson DM, Eisenman JR, Fung V, Grabstein KH;
XX PI Rauch C;
XX WPI: 1999-253930/21.
XX DR N-PSDB; X29479.
XX PT Antibodies specific for epithelium-derived T-cell growth factor
XX PS Claim 1; Fig 1; 34pp; English.

CC The invention relates to an isolated antibody that binds specifically to
CC a simian or human epithelium-derived T-cell factor (ETF) polypeptide. The
CC antibodies are used, optionally when immobilized or labeled, to detect
CC and quantify ETF in standard immunoassays. They may also be used as
CC diagnostic and therapeutic agents, e.g. when conjugated to toxins (or
CC their precursors) or radionuclides. ETF induces proliferation and/or

CC differentiation of T cells (or their precursors), e.g. for use in
CC establishing long term in vitro cultures; and is also used to treat
CC gastrointestinal disease (e.g. enteritis or mucositis) induced by
CC chemotherapy or radiation, peptic ulcer, gastroenteritis, colitis,
CC villus atrophy, malignancy and inflammatory bowel disease), to treat
CC human immune deficiency virus infection or associated disease, or
CC generally in any situation requiring stimulation of T or B cell
CC proliferation, secretion of immunoglobulins or certain cytokines,
CC increased anti-infectious disease immunity, induction of T-cell lytic
CC activity or increased destruction of tumour or virus-infected cells. The
CC present sequence represents the mature active sequence of simian ETF
XX polypeptide.

XX SQ Sequence 114 AA;

Query Match 14.3%; Score 99.5; DB 20; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00062;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSEAFSCF---QKAQLKSANTGNNE 62
Db 3 vnvvisdlkk-iedliqsmhidatlytesdvhpckvtamkcfllleqlvislesgasihd 61
QY 63 RIINVSIIKKLRKPPSTNAGRQKRLTCPCDSYKPKPEFLERFKSLLOKMIH 118
Db 62 tvenlii--lannslsngnvtes---gckeceeleeknikelfqsfvhiqvmfin 112

RESULT 15
Y52309
ID Y52309 standard; Protein; 114 AA.
XX AC Y52309;
XX DT 09-FEB-2000 (first entry)
XX DE Mature simian epithelium-derived T-cell factor (ETF).
XX KW ETF; epithelium-derived T-cell factor; T-cell; T-lymphocyte;
KW proliferation; differentiation; growth factor; precursor; mature; CD4+;
KW CD8+; gastrointestinal; epithelium; cell culture; treatment; HIV;
KW gastrointestinal disease; gastroenteritis; colitis;
KW inflammatory bowel disease; villus atrophic disorder; enteritis;
KW chemotherapy; radiotherapy; gut toxicity; cancer; side effect;
XX OS Mammalia.
XX PN US5985262-A.
XX PD 16-NOV-1999.
XX PF 03-FEB-1997; 97US-0794524.
XX PR 22-FEB-1995; 95US-0393305.
XX PR 04-OCT-1996; 96US-0726817.
XX PR 08-MAR-1993; 93US-0031399.
XX PR 22-APR-1994; 94US-0233606.
XX PA (IMMV) IMMUNEX CORP.
XX PI Rauch C, Fung V, Eisenman JR, Grabstein KH, Anderson DM;
XX DR WPI: 2000-022267/02.
XX DR N-PSDB; Z38244.

CC Stimulation of T-cells in human immunodeficiency virus infected
CC patients -
XX Claim 1; Fig 1; 33pp; English.
XX This sequence represents mature simian epithelium-derived T-cell factor

(ETF). This is a previously unidentified T-cell growth factor which stimulates precursor or mature CD4+ and CD8+ T-lymphocytes to proliferate and differentiate. It also promotes proliferation of the gastrointestinal epithelium. The protein can be used to promote long-term in vitro culture of T-lymphocytes and T-cell lines. ETF can be used for treating HIV infection, HIV-associated diseases, and other diseases or conditions where stimulation of T-cell proliferation would be desirable e.g., it could be used to augment the destruction of tumour cells or virally-infected cells. ETF may also be used to treat or prevent gastrointestinal disease, including chemotherapy and radiotherapy associated enteritis, gastroenteritis, colitis, inflammatory bowel disease and villus atrophic disorders. Chemotherapy and radiotherapy associated enteritis (gut toxicity) results in bleeding and sepsis due to gastrointestinal flora entering the blood, and thus can limit the dosage of therapeutic agent administered to a cancer patient. ETF may therefore be used to increase the tolerated doses radiotherapy and chemotherapy.

AA	Sequence	114-AA;
SQ		

Query Match	14.3%	Score 99.5	DB 21	Length 114
Best Local Similarity	25.9%	Pred. No. 0.00062		
Matches 30	Conservative	23	Mismatches 48	Indels 15
				Gaps 5

Qy 12 IDIVDLKYNVNDLVPEF-----LPAPEDVETNCEWSAFSCF-----QKAQLKSANTGNNE 62

Db 3 vnvisdlkk-iedliqsmhidatlytstdvhpksckvtamkcflllelqlvislesgdasihd 61

```

63 RIINVSIIKKLRKPPSTNAGRROKHRITCPCDSYEEKPPKEFLERFKSLQKMIH 118
   : : | | | | : | : | | | | : : | :
62 tvenlii--lannslsngntes---gckeeceeleeknikelfqsfvhiqvmfin 112

```

Search completed: May 23, 2001, 11:11:34
Job time: 177 sec

Result No.	Query			Length	DB	ID	Description
	Score	Match					
1	99.5	14.3	114	1	US-08-031-399-6	Sequence 6, Appli	
2	99.5	14.3	114	1	US-08-031-399-12	Sequence 12, Appl	
3	99.5	14.3	114	1	US-08-393-305-3	Sequence 3, Appli	
4	99.5	14.3	114	1	US-08-726-817-3	Sequence 3, Appli	
5	99.5	14.3	114	1	US-08-504-042-6	Sequence 6, Appli	
6	99.5	14.3	114	1	US-08-504-042-12	Sequence 12, Appli	
7	99.5	14.3	114	2	US-08-725-969-3	Sequence 3, Appli	
8	99.5	14.3	114	2	US-08-794-524-3	Sequence 3, Appli	
9	99.5	14.3	114	4	US-09-189-193-3	Sequence 3, Appli	
10	99.5	14.3	114	5	PCF-US94-03793-6	Sequence 6, Appli	
11	99.5	14.3	114	5	PCF-US94-03793-12	Sequence 12, Appli	
12	99.5	14.3	162	1	US-08-031-399-5	Sequence 5, Appli	
13	99.5	14.3	162	1	US-08-393-305-2	Sequence 2, Appli	
14	99.5	14.3	162	1	US-08-535-733-2	Sequence 2, Appli	
15	99.5	14.3	162	1	US-08-726-817-2	Sequence 2, Appli	
16	99.5	14.3	162	1	US-08-504-042-5	Sequence 5, Appli	
17	99.5	14.3	162	2	US-08-725-969-2	Sequence 2, Appli	
18	99.5	14.3	162	2	US-08-794-524-2	Sequence 2, Appli	
19	99.5	14.3	162	3	US-08-842-947-6	Sequence 6, Appli	
20	99.5	14.3	162	4	US-09-189-193-2	Sequence 2, Appli	
21	99.5	14.3	162	5	PCF-US94-03793-5	Sequence 5, Appli	
22	99.5	14.3	162	5	PCF-US96-06423-2	Sequence 2, Appli	
23	94.5	13.6	114	1	US-08-031-399-3	Sequence 3, Appli	
24	94.5	13.6	114	1	US-08-393-305-6	Sequence 6, Appli	
25	94.5	13.6	114	1	US-08-726-817-6	Sequence 6, Appli	
26	94.5	13.6	114	1	US-08-504-042-3	Sequence 3, Appli	
27	94.5	13.6	114	2	US-08-725-969-6	Sequence 6, Appli	

Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTAMKCFLEQLQVISLQSGDASIHD 61
QY 63 RIINVSIRKKLKKRPPSTNAGRQKRLTCPCSDSYEKPPKPEFLERFKSLQKMIH 118
Db 62 TVENLII--LANNSSNGNVTES---GCKECELEEKNIKEFLQSFVHIVQMFN 112

RESULT 2
US-08-031-399-12
; Sequence 12, Application US/08031399
; Patent No. 5552303
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Epithelium-derived T-cell Factor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/031.399
; FILING DATE: 19930308
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-031-399-12

Query Match 14.3%; Score 99.5; DB 1; Length 114;
Best Local Similarity 25.2%; Pred. No. 0.00031;
Matches 30; Conservative 23; Mismatches 45; Indels 21; Gaps 5;
QY 12 IDIVDLKKNYNDLVPEF-----LPAPEDVTNCWESAFSCF-QKAOLKSANTGN--- 61
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTAMKCFLEQLQVISLQSGDXXIHD 61
QY 62 --ERINVSIRKKLKKRPPSTNAGRQKRLTCPCSDSYEKPPKPEFLERFKSLQKMIH 118
Db 62 TVENLIIANNLSSNGNVTESG-----CKECELEEKNIKEFLQSFVHIVQMFN 112

RESULT 3
US-08-393-305-3
; Sequence 3, Application US/08393305
; Patent No. 5574138
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor

; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,305
; FILING DATE: 22-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-393-305-3

Query Match 14.3%; Score 99.5; DB 1; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00031;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;
QY 12 IDIVDLKKNYNDLVPEF-----LPAPEDVTNCWESAFSCF---QKAOLKSANTGNNE 62
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTAMKCFLEQLQVISLQSGDASIHD 61
QY 63 RIINVSIRKKLKKRPPSTNAGRQKRLTCPCSDSYEKPPKPEFLERFKSLQKMIH 118
Db 62 TVENLII--LANNSSNGNVTES---GCKECELEEKNIKEFLQSFVHIVQMFN 112

RESULT 4
US-08-726-817-3
; Sequence 3, Application US/08726817
; Patent No. 5707616
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,817


```
Query Match      14.3%   Score 99.5; DB 1; Length 114;  
Best Local Similarity 25.2%; Pred. No. 0.00031;  
Matches 30; Conservative 23; Mismatches 45; Indels 21; Gaps 5;  
  
y          12 IDIVDQIKYVNDLVPEF-----LPAPEDVTETNCWSAFSCF-QKAQLKSANTGNN----- 61  
           ::::: || | : ||:  
           :
```

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/189,193
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/393,305
FILING DATE: 22-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: McMasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 480052.409C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 114 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-189-193-3

Query Match 14.3%; Score 99.5; DB 4; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00031;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;
QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKAQLKSANTGNNE 62
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTSDVHPSCVKVTAMKCFLLLEQLQVISLSDGASIH 61
QY 63 RIINVSIIKKLRKPPSTNAGRQKHRLTQPCSDSYEKKPPKFLERFKSLLOKMIH 118
Db 62 TVENLII--LANNLSNGNVTES---GCKECELEEKNIKEFLQSFVHVQMF 112

RESULT 10
PCT-US94-03793-6
Sequence 6, Application PC/TUS9403793
GENERAL INFORMATION:
APPLICANT: Grabstein, Kenneth
APPLICANT: Anderson, Dirk
APPLICANT: Eisenman, June
APPLICANT: Fung, Victor
APPLICANT: Rauch, Charles
TITLE OF INVENTION: Interleukin-15
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/03793
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Launer, Charlene
REGISTRATION NUMBER: 33,035
REFERENCE/DOCKET NUMBER: 2811
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-587-0430
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:

LENGTH: 114 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
PCT-US94-03793-6

Query Match 14.3%; Score 99.5; DB 5; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00031;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;
QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKAQLKSANTGNNE 62
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTSDVHPSCVKVTAMKCFLLLEQLQVISLSDGASIH 61
QY 63 RIINVSIIKKLRKPPSTNAGRQKHRLTQPCSDSYEKKPPKFLERFKSLLOKMIH 118
Db 62 TVENLII--LANNLSNGNVTES---GCKECELEEKNIKEFLQSFVHVQMF 112

RESULT 11
PCT-US94-03793-12
Sequence 12, Application PC/TUS9403793
GENERAL INFORMATION:
APPLICANT: Grabstein, Kenneth
APPLICANT: Anderson, Dirk
APPLICANT: Eisenman, June
APPLICANT: Fung, Victor
APPLICANT: Rauch, Charles
TITLE OF INVENTION: Interleukin-15
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/03793
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Launer, Charlene
REGISTRATION NUMBER: 33,035
REFERENCE/DOCKET NUMBER: 2811
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-587-0430
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 114 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
PCT-US94-03793-12

Query Match 14.3%; Score 99.5; DB 5; Length 114;
Best Local Similarity 25.2%; Pred. No. 0.00031;
Matches 30; Conservative 23; Mismatches 45; Indels 21; Gaps 5;
QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-QKAQLKSANTGN-- 61
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTSDVHPSCVKVTAMKCFLLLEQLQVISLSDGASIH 61
QY 62 --ERIINVSIIKKLRKPPSTNAGRQKHRLTQPCSDSYEKKPPKFLERFKSLLOKMIH 118

Db 62 TVENLIIANNXLSNGNXTESG-----CKECELEEKNIKEFLQSFVHVQMFN 112

RESULT 12

US-08-031-399-5

Sequence 5, Application US/08031399

Patent No. 5552303

GENERAL INFORMATION:

APPLICANT: Grabstein, Kenneth

APPLICANT: Anderson, Dirk

APPLICANT: Eisenman, June

APPLICANT: Fung, Victor

APPLICANT: Rauch, Charles

TITLE OF INVENTION: Epithelium-derived T-cell Factor

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Immunex Corporation

STREET: 51 University Street

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/031.399

FILING DATE: 19930308

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Launer, Charlene

REGISTRATION NUMBER: 33,035

REFERENCE/DOCKET NUMBER: 2811

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-587-0430

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 162 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-031-399-5

Query Match 14.3%; Score 99.5; DB 1; Length 162;

Best Local Similarity 25.9%; Pred. No. 0.00048;

Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKAQLKSANTGNNE 62

Db 51 VNVISDLAK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVLSLESGDASIH 109

QY 63 RIINVSIIKKLRKPPSTNAGRQKRLTCPCDSYVEKKPKPEFLERFKSLQKMH 118

Db 110 TVENLII--LANNLSNGNVTES---GCKECELEEKNIKEFLQSFVHVQMFN 160

RESULT 13

US-08-393-305-2

Sequence 2, Application US/08393305

Patent No. 5574138

GENERAL INFORMATION:

APPLICANT: Grabstein, Kenneth

APPLICANT: Anderson, Dirk

APPLICANT: Eisenman, June

APPLICANT: Fung, Victor

APPLICANT: Rauch, Charles

TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/393.305

FILING DATE: 22-FEB-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: McMasters, David D.

REGISTRATION NUMBER: 33,963

REFERENCE/DOCKET NUMBER: 480052.409C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-622-4900

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 162 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-393-305-2

Query Match 14.3%; Score 99.5; DB 1; Length 162;

Best Local Similarity 25.9%; Pred. No. 0.00048;

Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKAQLKSANTGNNE 62

Db 51 VNVISDLAK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVLSLESGDASIH 109

QY 63 RIINVSIIKKLRKPPSTNAGRQKRLTCPCDSYVEKKPKPEFLERFKSLQKMH 118

Db 110 TVENLII--LANNLSNGNVTES---GCKECELEEKNIKEFLQSFVHVQMFN 160

RESULT 14

US-08-535-733-2

Sequence 2, Application US/08535733

Patent No. 5660824

GENERAL INFORMATION:

APPLICANT: Quinn, LeBris

APPLICANT: Grabstein, Kenneth

APPLICANT: Troutt, Anthony B.

TITLE OF INVENTION: Muscle-Trophic Factor

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunex Corporation

STREET: 51 University Street

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Apple Macintosh

OPERATING SYSTEM: System 7, Word 5.1a

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/535.733

FILING DATE:

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Malaska, Stephen L.

REGISTRATION NUMBER: 32,655

REFERENCE/DOCKET NUMBER: 2833

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-587-0430

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 162 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-535-733-2

Query Match

14.3%; Score 99.5; DB 1; Length 162;

Best Local Similarity 25.9%; Pred. No. 0.00048;

Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-----OKAOLKSANTGNNE 62
Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLEQLQVISLESGLDASIH 109
QY 63 RIINVSIKKLRKPPSTNAGRQKHRLTCPCSDSYEKKPPKPEFLERFKSLLOKMIH 118
Db 110 TVENLII--LANNSSLSNGNVTES---GCKECELEEKNIKEFLQSFVHIVQMF 160

RESULT 15

US-08-726-817-2

Sequence 2, Application US/08726817

Patent No. 5707616

GENERAL INFORMATION:

APPLICANT: Grabstein, Kenneth

APPLICANT: Anderson, Dirk

APPLICANT: Eisenman, June

APPLICANT: Fung, Victor

APPLICANT: Rauch, Charles

TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/726,817

FILING DATE: 04-OCT-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/393,305

FILING DATE: 22-FEB-1995

ATTORNEY/AGENT INFORMATION:

NAME: Mcmasters, David D.

REGISTRATION NUMBER: 33,963

REFERENCE/DOCKET NUMBER: 480052.409C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-622-4900

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 162 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-726-817-2

Query Match

14.3%; Score 99.5; DB 1; Length 162;

Best Local Similarity 25.9%; Pred. No. 0.00048;

Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-----OKAOLKSANTGNNE 62
Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLEQLQVISLESGLDASIH 109
QY 63 RIINVSIKKLRKPPSTNAGRQKHRLTCPCSDSYEKKPPKPEFLERFKSLLOKMIH 118
Db 110 TVENLII--LANNSSLSNGNVTES---GCKECELEEKNIKEFLQSFVHIVQMF 160

Search completed: May 23, 2001, 11:09:41

Job time: 64 sec

Result No.	Score	Query %		Length	DB	ID	Description
		Match	Mismatch				
1	94.5	13	6	162	1	A53484	interleukin-15 pre
2	91.5	13.2		567	2	S29498	lymphocyte antigen
3	82	11.8		607	1	ABXL72	74K albumin precu
4	79	11.4		206	2	S49882	hypothetical prote
5	75.5	10.9		336	2	S43632	Fit-1S protein pre
6	74	10.6		440	2	T47906	FUSCA PROTEIN FUS6
7	74	10.6		2197	2	B71600	variant-specific s
8	73.5	10.6		162	2	I49124	interleukin-15 - m
9	73.5	10.6		472	2	IS9087	ISG-K54 - human
10	73.5	10.6		868	2	T31527	hypothetical prote
11	73	10.5		304	2	A32108	translation initia
12	73	10.5		848	2	T00372	hypothetical prote
13	72.5	10.4		309	2	T41889	pe38 or fl53 - Bomb
14	72.5	10.4		1101	2	T21819	hypothetical prote
15	72.5	10.4		1206	2	S24407	formin isoform IV
16	72	10.4		257	2	G69230	conserved hypothet
17	72	10.4		262	2	F73858	probable methyl tr
18	72	10.4		262	2	T41813	ACMNPV orf69 - Bom
19	71.5	10.3		195	2	S42022	ureidoglycolate hy
20	71.5	10.3		741	2	A83317	isocitrate dehydro
21	71	10.2		321	2	T46352	hypothetical prote
22	71	10.2		406	2	T28957	hypothetical prote
23	71	10.2		719	2	T27977	lin-15A protein -
24	71	10.2		805	2	A46266	aryl hydrocarbon r
25	71	10.2		864	2	T08575	protein kinase hom
26	71	10.2		2206	2	G71611	hypothetical prote
27	71	10.2		2324	2	T18378	variant-specific s
28	70.5	10.1		447	2	S39316	CAB3b protein - hu
29	70.5	10.1		477	2	S21049	calcium channel pr

S29498
Lymphocyte antigen Ly84 precursor
N:Alternate names: 38.5K T1 glycoprotein; ST2L protein
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 01-Dec-2000
C:Accession: S29498; A33541; S17657; S07054
R:Yanagisawa, K.; Takagi, T.; Tsukamoto, T.; Tetsuka, T.; Tominaga, S.
FEBS Lett. 318, 83-87, 1993
A:Title: Presence of a novel primary response gene ST2L, encoding a product
A:Reference number: S29498; MUID: 93170492
A:Accession: S29498
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-567 <YAN>
A:Cross-references: EMBL:D13695; NID:q286100; PID:q286101

R:Klemenz, R.; Hoffmann, S.; Werenskiold, A.K.
 Proc. Natl. Acad. Sci. U.S.A. 86, 5708-5712, 1989
 A:Title: Serum- and oncoprotein-mediated induction of a gene with sequence similarity to
 A:Reference number: A33541; MUID:89345536
 A:Accession: A33541
 A:Molecule type: mRNA
 A:Residues: 1-191, 'V', 193-328, 'SKECPSHIA' <KLE>
 A:Cross-references: GB:M24843; NID:g201103; PIDN:AAA40160.1; PID:g201104
 R:Tominaga, S.I.; Jenkins, N.A.; Gilbert, D.J.; Copeland, N.G.; Tetsuka, T.
 Biochim. Biophys. Acta 1090, 1-8, 1991
 A:Title: Molecular cloning of the murine ST2 gene. Characterization and chromosomal map
 A:Reference number: S17657; MUID:91355215
 A:Accession: S17657
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-191, 'V', 193-328, 'SKECPSHIA' <TOM>
 A:Cross-references: EMBL:X60184; NID:g54200; PIDN:CAA42742.1; PID:g54201
 R:Tominaga, S.I.
 FEBS Lett. 258, 301-304, 1989
 A:Title: A putative protein of a growth specific cDNA from BALB/c-3T3 cells is highly si
 A:Reference number: S07034; MUID:90092495
 A:Accession: S07034
 A:Molecule type: mRNA
 A:Residues: 1-328, 'SKECPSHIA' <FO2>
 A:Cross-references: EMBL:Y07519; NID:g55517; PIDN:CAA68812.1; PID:g55518
 A:Note: it is uncertain whether Met-1, Met-7 or Met-19 is the initiator
 C:Genetics:
 A:Gene: ST2
 A:Map position: 1
 A:Introns: 27/1; 97/2; 155/3; 210/1; 233/1; 280/2
 C:Superfamily: interleukin-1 receptor type I
 C:Keywords: glycoprotein
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-337/Product: ST2 protein #status predicted <MAT>
 F:60,101,107,146,176,194,225,259,278/Binding site: carbohydrate (Asn) (covalent) #status

Query Match 13.2%; Score 91.5; DB 2; Length 567;
 Best Local Similarity 25.6%; Pred. No. 0.45;
 Matches 33; Conservative 20; Mismatches 39; Indels 37; Gaps 7;
 QY 16 DLKKNVNDLVPEFLPAPEDVETNCWSAFSCFQAKLKSANTGNININYSIKLKRK 75
 Db 76 DRLK-----FLPARVE-----DSGIYACV-----IRSNPL-NKTGYLVNTIHK---K 113
 QY 76 PPSTN-----AGRRQKRLTGPCSDSYEKPPKEFLERFKSLQKMIHQH----- 120
 Db 114 PPSNIPDLYMYSTVRGSDKNKICTPTIDLYNTAPVQWFKNCKALQEPFRHRSYLF 173
 QY 121 LSSRTHGSE 129
 Db 174 IDNVTHDDE 182
 RESULT 3
 ABXL72
 74k albumin precursor - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 23-Jun-1999
 C:Accession: B41682; S02693; A05288
 R:Moskaitis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.
 Mol. Endocrinol. 5, 464-473, 1989
 A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic ac
 e during development.
 A:Reference number: A41682; MUID:89313788
 A:Accession: B41682
 A:Molecule type: mRNA
 A:Residues: 3-607 <MOS>
 A:Cross-references: GB:M21442; NID:g213930; PIDN:AAA49637.1; PID:g213931
 R:Schorpp, M.; Doebebling, U.; Wagner, U.; Ryffel, G.U.
 J. Mol. Biol. 199, 83-93, 1988
 A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Dele
 A:Reference number: S02692; MUID:88172470

A:Accession: S02693
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-48 <SCH>
 A:Cross-references: EMBL:Z26826
 R:Wolffe, A.P.; Glover, J.F.; Martin, S.C.; Tenniswood, M.P.R.; Williams, J.L.; Tata,
 Eur. J. Biochem. 146, 489-496, 1985
 A:Title: Deinduction of transcription of Xenopus 74-kDa albumin genes and destabiliza
 A:Reference number: A05288; MUID:85126974
 A:Accession: A05288
 A:Molecule type: mRNA
 A:Residues: 459-502, 'L', 504-557 <WOL>
 A:Cross-references: GB:M28276
 A:Note: the authors translated the codon TAT for residue 63 as Thr
 C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds coppe
 mones (weak bonds with these hormones promote their transfer across the membranes), t
 C:Genetics:
 A:Introns: 27/1
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; glycoprotein; metal binding; plasma
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-24/Domain: propeptide #status predicted <PRO>
 F:25-607/Product: 74k serum albumin #status predicted <MAT>
 F:32-201/Domain: serum albumin repeat homology <SA1>
 F:220-393/Domain: serum albumin repeat homology <SA2>
 F:412-591/Domain: serum albumin repeat homology <SA3>
 F:30/Binding site: copper (His) #status predicted
 F:80-88,101-117,116-127,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-3
 F:256/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 11.8%; Score 82; DB 1; Length 607;
 Best Local Similarity 24.2%; Pred. No. 3.9;
 Matches 38; Conservative 19; Mismatches 50; Indels 50; Gaps 6;
 QY 13 DIVDQLKNYNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKALQKSN 57
 Db 72 EINDFAKSCINDKTPECEKPVGTLLFDKLCADPAVGVMYEWSEKCAKQDPERACQCFKAH 131
 QY 58 TGNNERII-----NVSIKLKRK-----PPSTNAGRQKRLTCTPS 93
 Db 132 RDHEHTSIKPEPEETCKLKEHPDILLSAFIEARNHPDLYPPAVLALTQYHKLAHC 191
 QY 94 CDSYEKKPKPEFLERFKSLQKMIHQHLSSTHGSED 130
 Db 192 CEEDKE--KCFSEKMKQLMK-----QSHSIED 217
 RESULT 4
 S49882
 hypothetical protein Y1127c - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein Y18277.O2c
 C:Species: Saccharomyces cerevisiae
 C:Date: 13-Jan-1995 #sequence_revision 20-Feb-1995 #text_change 05-Nov-1999
 C:Accession: S49882
 R:Hamlyn, N.; Churcher, C.
 submitted to the EMBL Data Library, November 1994
 A:Reference number: S49881
 A:Accession: S49882
 A:Molecule type: DNA
 A:Residues: 1-206 <HAM>
 A:Cross-references: GB:Z47047; EMBL:Z46833; NID:g603997; PID:g763219; GSPDB:GN00009;
 C:Genetics:
 A:Gene: MIPS:Y1127C
 A:Map position: 9L
 Query Match 11.4%; Score 79; DB 2; Length 206;
 Best Local Similarity 26.1%; Pred. No. 2.4;
 Matches 29; Conservative 18; Mismatches 48; Indels 16; Gaps 3;
 QY 17 QLNKYNVDLVPEFLPAPEDVETNCWSAFSCFQAKLKSANTGNININYSIKLKRK 76
 Db 174 IDNVTHDDE 182


```

Db      12 QATSVVNGLSNLPLPGVPKIRANNGKTSVNNNGSKAQLIDRNLRKRVQLQNRDVIHKIKKC 71
QY      77 PSTNAGRQRQKHLRTPCSDSYEKKPKPKFELEF--KSLLOKMIHQHLSRFT 125
Db      72 KLVRKKVKVKKHL-----DKLEQLAKHQVLKK--HQEECTLT 108

RESULT 5
S42632
Fit-15 protein precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 25-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 19-May-2000
C:Accession: S42632
R:Bergers, G.; Reikerstorfer, A.; Braselmann, S.; Graninger, P.; Busslinger, M.
EMBO J. 13, 1176-1188, 1994
A:Title: Alternative promoter usage of the Fos-responsive gene Fit-1 generates mRNA iso
A:Reference number: S42632; MUID:94178260
A:Accession: S42632
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-336 <BR>
A:Cross-references: GB:U04319; NID:g488278; PIDN:AAA67172.1; PID:g488279
C:Superfamily: vaccinia virus B15R protein; immunoglobulin homology

Query Match          10.9%; Score 75.5; DB 2; Length 336;
Best Local Similarity 22.2%; Pred. No. 8.6;
Matches 30; Conservative 26; Mismatches 42; Indels 37; Gaps 6;

QY      12 IDIVDLKNVNDLVP-----EFLPAPEDVETNCWSAFSCFQAQLKSAN 57
      1: |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db      49 INPVEWYTSNTNRIPTQKRNIQVSRDLKFLPAKVE-----DSGIYTCVIRSP-ESIK 102
      1::|::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY      58 TGNNERIINYSIKLKRKPP-----STNAGRQKHRLTCPCSDSYEKKPKPEFLE 107
      1::|::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db      103 TGS-----LNTVIYK---RPPNCKIPDYMYSTVDGSDKNKITCTPTIALYNWTAPVQWFK 155
      1::|::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY      108 RFKSLLOKMIHQHLS 122
      1::|::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db      156 NCKALQGRFRAHMS 170
      1::|::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|

RESULT 6
T47906
FUSCA PROTEIN FUS6 - Arabidopsis thaliana
N:Alternate names: protein T20K12.40
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47906
R:De Haan, M.; Maarse, A.C.; Grivell, L.A.; Meves, H.W.; Lemcke, K.; Mayer, K.F.X.; Que
Submitted to the Protein Sequence Database, January 2000
A:Reference number: 224480
A:Accession: T47906
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-440 <DEH>
A:Cross-references: EMBL:AL137898
A:Experimental source: cultivar Columbia; BAC clone T20K12
C:Genetics:

Query Match          10.6%; Score 74; DB 2; Length 440;
Best Local Similarity 22.0%; Pred. No. 16;
Matches 26; Conservative 22; Mismatches 50; Indels 20; Gaps 3;

QY      3 RHMTMRQLDIVDLKNVNDLVPFLPAPEDVETNCWSAFSCFQAQLKSANTGNNE 62
      ::|::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db      230 KYIAARKFLDVPNGLSNVNEVI-----APQDIATYGGICALASFDRSELKAF----- 278
      ::|::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY      63 RIINVSIKLKRKPPSTNAGRQKHRLTCPCSDSYEKKPKPKFELEFRLFKSLLOKMIHQH 120

```

[illegible]

```

I59087
ISG-K54 - human
C:Species: Homo sapiens (man)
C>Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 29-Sep-1999
C:Accession: I59087
R:Levy, D.; Larner, A.; Chaudhuri, A.; Babiss, L.E.; Darnell, J.E.
Proc. Natl. Acad. Sci. U.S.A. 83, 8929-8933, 1986
A>Title: Interferon-stimulated transcription: isolation of an inducible gene and identification of its protein product
A:Reference number: I59087; MUID:87067427
A:Accession: I59087
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-472 <RES>
A:CROSS-references: GB:M14660; NID:g186559; PIDN:AAA59191.1; PID:g386836
C:Genetics:
A:Gene: GDB:IFI54; GI0P2; IFI-54
A:CROSS-references: GDB:I19959; OMIM:147040
A:Map position: 10q23-10q24
A:Introns: 2/2
C:Superfamily: interferon-induced 56k protein

Query Match      10.6%; Score 73.5; DB 2; Length 472;
Best Local Similarity 26.9%; Pred. No.19;
Matches 39; Conservative 18; Mismatches 49; Indels 39; Gaps 8;

QY    4 HMIRMRLQIDIVDLKKNVNDLVPEFLPAP---EDVTNCE--WS-----AFSC 47
       ||| : |::||: | | | | | | | | | | | | | | | | | | | | |
Db     106 HMGRLSDVIYDKVKH-----VCEFSPRYTESPELDCGEWTRLKCGGNQNERAKVC 160
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY    48 FOKAQLSANTGNRRINIVSIIKKLRKPSTNA--GRROKHRLTCPSDSYEKKPKPEF 105
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     161 FEKALEKKPKNPFTSGLAIASYRLDWWPPSQNAIDLROAIRL-----NPDNQY 210
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY    106 LERFKSLILQKIHQHLSSRTGISED 130
       | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     211 L---KVLLALKLHK---MREEGEEE 229
       | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 10
T31527
hypothetical protein Y47D3A.14 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T31527
R:Matthews, L.
submitted to the EMBL Data Library, October 1999
A:Reference number: Z21043
A:Accession: T31527
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-868 <WIL>
A:CROSS-references: EMBL:AL117202; PIDN:CAB55073.1; CESP:Y47D3A.14
A:Experimental source: clone Y47D3A
C:Genetics:
A:Gene: CESP:Y47D3A.14
A:Introns: 17/2; 51/3; 102/3; 169/3; 246/1; 287/2; 319/3; 371/3; 395/2; 440/3; 450/3; 499/3; 500/3; 501/3; 502/3; 503/3; 504/3; 505/3; 506/3; 507/3; 508/3; 509/3; 510/3; 511/3; 512/3; 513/3; 514/3; 515/3; 516/3; 517/3; 518/3; 519/3; 520/3; 521/3; 522/3; 523/3; 524/3; 525/3; 526/3; 527/3; 528/3; 529/3; 530/3; 531/3; 532/3; 533/3; 534/3; 535/3; 536/3; 537/3; 538/3; 539/3; 540/3; 541/3; 542/3; 543/3; 544/3; 545/3; 546/3; 547/3; 548/3; 549/3; 550/3; 551/3; 552/3; 553/3; 554/3; 555/3; 556/3; 557/3; 558/3; 559/3; 560/3; 561/3; 562/3; 563/3; 564/3; 565/3; 566/3; 567/3; 568/3; 569/3; 570/3; 571/3; 572/3; 573/3; 574/3; 575/3; 576/3; 577/3; 578/3; 579/3; 580/3; 581/3; 582/3; 583/3; 584/3; 585/3; 586/3; 587/3; 588/3; 589/3; 590/3; 591/3; 592/3; 593/3; 594/3; 595/3; 596/3; 597/3; 598/3; 599/3; 600/3; 601/3; 602/3; 603/3; 604/3; 605/3; 606/3; 607/3; 608/3; 609/3; 610/3; 611/3; 612/3; 613/3; 614/3; 615/3; 616/3; 617/3; 618/3; 619/3; 620/3; 621/3; 622/3; 623/3; 624/3; 625/3; 626/3; 627/3; 628/3; 629/3; 630/3; 631/3; 632/3; 633/3; 634/3; 635/3; 636/3; 637/3; 638/3; 639/3; 640/3; 641/3; 642/3; 643/3; 644/3; 645/3; 646/3; 647/3; 648/3; 649/3; 650/3; 651/3; 652/3; 653/3; 654/3; 655/3; 656/3; 657/3; 658/3; 659/3; 660/3; 661/3; 662/3; 663/3; 664/3; 665/3; 666/3; 667/3; 668/3; 669/3; 670/3; 671/3; 672/3; 673/3; 674/3; 675/3; 676/3; 677/3; 678/3; 679/3; 680/3; 681/3; 682/3; 683/3; 684/3; 685/3; 686/3; 687/3; 688/3; 689/3; 690/3; 691/3; 692/3; 693/3; 694/3; 695/3; 696/3; 697/3; 698/3; 699/3; 700/3; 701/3; 702/3; 703/3; 704/3; 705/3; 706/3; 707/3; 708/3; 709/3; 710/3; 711/3; 712/3; 713/3; 714/3; 715/3; 716/3; 717/3; 718/3; 719/3; 720/3; 721/3; 722/3; 723/3; 724/3; 725/3; 726/3; 727/3; 728/3; 729/3; 730/3; 731/3; 732/3; 733/3; 734/3; 735/3; 736/3; 737/3; 738/3; 739/3; 740/3; 741/3; 742/3; 743/3; 744/3; 745/3; 746/3; 747/3; 748/3; 749/3; 750/3; 751/3; 752/3; 753/3; 754/3; 755/3; 756/3; 757/3; 758/3; 759/3; 760/3; 761/3; 762/3; 763/3; 764/3; 765/3; 766/3; 767/3; 768/3; 769/3; 770/3; 771/3; 772/3; 773/3; 774/3; 775/3; 776/3; 777/3; 778/3; 779/3; 780/3; 781/3; 782/3; 783/3; 784/3; 785/3; 786/3; 787/3; 788/3; 789/3; 790/3; 791/3; 792/3; 793/3; 794/3; 795/3; 796/3; 797/3; 798/3; 799/3; 800/3; 801/3; 802/3; 803/3; 804/3; 805/3; 806/3; 807/3; 808/3; 809/3; 810/3; 811/3; 812/3; 813/3; 814/3; 815/3; 816/3; 817/3; 818/3; 819/3; 820/3; 821/3; 822/3; 823/3; 824/3; 825/3; 826/3; 827/3; 828/3; 829/3; 830/3; 831/3; 832/3; 833/3; 834/3; 835/3; 836/3; 837/3; 838/3; 839/3; 840/3; 841/3; 842/3; 843/3; 844/3; 845/3; 846/3; 847/3; 848/3; 849/3; 850/3; 851/3; 852/3; 853/3; 854/3; 855/3; 856/3; 857/3; 858/3; 859/3; 860/3; 861/3; 862/3; 863/3; 864/3; 865/3; 866/3; 867/3; 868/3; 869/3; 870/3; 871/3; 872/3; 873/3; 874/3; 875/3; 876/3; 877/3; 878/3; 879/3; 880/3; 881/3; 882/3; 883/3; 884/3; 885/3; 886/3; 887/3; 888/3; 889/3; 890/3; 891/3; 892/3; 893/3; 894/3; 895/3; 896/3; 897/3; 898/3; 899/3; 900/3; 901/3; 902/3; 903/3; 904/3; 905/3; 906/3; 907/3; 908/3; 909/3; 910/3; 911/3; 912/3; 913/3; 914/3;
```

A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 'E',122-81
A;Cross-references: EMBL
A;Experimental source
C;Genetics:
A;Note: KIAA0650; DKFZ

	Query Match	10.5%	Score 73;	DB 2;	Length 848;
	Best Local Similarity	25.5%;	Pred. No. 39;		
	Matches	27;	Conservative	22;	Mismatches 47; Indels 10; Gaps 5;
Oy	28 EFLPAPDEVETNC-EWSAFSCFQAQKLSANTGNNERLIINYSIKLKKRPSTNAGRQK	86	: : :	:	:
Db	66 KFIPGPNGKLCFTWREFSDFIRVOLIS---GPPAKLLLDWPCLKESIPVIN-GRDLQ	121	: : :	:	:
Oy	87 HRLTCPDSCSYEKPPKEFLERFKSLQKMTHQLH--SSRTHGSED	130	: : :	:	:
Db	122 NPITIVOLCDOWNDNPAP---VOHVKSITLKASNLKLMPSPNOOHKTDE	164	: : :	:	:

```

RESULT 13
T41889
PE38 orf153 - Bombyx mori nuclear polyhedrosis virus (isolate T3)
C:Species: Bombyx mori nuclear polyhedrosis virus, BmSNPV
A:Variety: isolate T3
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: T41889
R:Gomi, S.; Majima, K.; Maeda, S.
J. Gen. Virol. 80, 1323-1337, 1999
A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
A:Reference number: 22020; MUID:99281911
A:Accession: T41889
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-309 <RAM>
A:Cross-references: EMBL:L33180; NID:g3745835; PIDN:AAC63818.1; PID:g3745971
A:Experimental source: isolate T3
C:Genetics:
A:Note: pe38
C:Superfamily: Orgyia pseudotsugata nuclear polyhedrosis virus PE-38 protein

```

```

Query Match      10.4%; Score 72.5; DB 2; Length 309;
Best Local Similarity 23.4%; Pred. No.15;
Matches 33; Conservative 30; Mismatches 57; Indels 21; Gaps 6;

Qy 7 RMROLIDIVDLKNYVNDLVPFLPAPEDVETNC-----EWSAFSCFQKAQIKSAN----- 57
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 148 KTOERVPAAVOYQKNIIVKLQERSVIVSENSONNCDDINNEOAKIVATAEALNKEKNHSDQ 207
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 58 -TGNNERII-----NVSIKKLRKPPSTNAGRQKHRLTCPSCDVSVEKKPKPFLEFRK 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 208 VTSENRLQIIEENTRLNEQVQLORQ-VRTLAPQR---GITVNPQIGRRDRAPAELENERFR 263

Qy 111 SLLQKMIHQ-HLSRTHGSED 130
   ||: | : : | : | :
Db 264 SLVSTISELFENRVHSIQN 284
   ||: | : : | : | :

```

RESULT 14
T26919
hypothetical protein Y45F10B.10 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26919
R:McMurray, A.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z20286
A:Accession: T26919
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA

A:Residues: 1-1101 <WIL>
A:Cross-references: EMBL:AL021487; PTDN:CAAL6357.1; GSPDB:GN00022; CESP:Y45F10B.10
A:Experimental source: Clone Y45F10B
C:Genetics:
A:Gene: CESP:Y45F10B.10
A:Map position: 4
A:Introns: 211/3; 292/1; 451/3; 556/2; 637/3; 755/3; 823/2; 946/1; 1078/3

	Query Match	10.4%	Score 72.5;	DB 2;	Length 1101;
	Best Local Similarity	31.5%;	Pred. No. 58;		
	Matches 23;	Conservative 14;	Mismatches 23;	Indels 13;	Gaps
Qy	49	KQAQLKSANTGNRIINVSIKLRRKP--PSTNAGRROK-HRLTPCSDSYEKKKPKPF	105		
		: : : : : : : : : : : : : : : :			
Db	1013	RSARAQSVSSANEPVASTSAGEIKKDPILSSNNGNAQSAPRATAP-----KPTFDM	1065		
		: : : : : : : : : : : : : : : :			
Qy	106	LERFK---SLIQK	115		
		: : : : : : : : : : : : : : : :			
Db	1066	LEPSKSRSTLSIEK	1078		
		: : : : : : : : : : : : : : : :			

RESULT 15
S24407
formin isoform IV - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C:Accession: S24407
R:Jackson-Grusby, L.; Kuo, A.; Leder, P.
Genes Dev. 6, 29-37, 1992
A:Title: A variant limb deformity transcript expressed in the embryonic mou
A:Reference number: S24407; MUID:92112033
A:Accession: S24407
A:Molecule type: mRNA
A:Residues: 1-1206 <JAC>
A:Cross-references: EMBL:X62379; NID:q51552; PIDN:CAA44244.1; PID:q51553

```

Query Match      10.4%; Score 72.5; DB 2; Length 1206;
Best Local Similarity 19.9%; Pred. No. 63;
Matches 32; Conservative 30; Mismatches 52; Indels 47; Gaps 6;

Qy      4  HMIMRQLIDIVDQLKNYVN-----DLVPEFLPAPEDV----- 36
      || : : : : || : || : || : ||
Db      960  HMKSVKDLAILAFNGYVMNGNTRGQADGYSLEILPKLKDVKSRONGMNLVDYVWKYY 1019
      || : : : : || : || : || : ||

Qy      37  -----ETNCWSAFSCFOKAQLKSANTGNNGRIINVSIKKLKPPKPPNAGRQKHRLT 90
      || : : : : || : || : || : ||
Db      1020  LRYDQEAGTDKSVFPLPEPODFVLASQVKEDLLK-LDRKLKQLEAS-----EQOMKLV 1074
      || : : : : || : || : || : ||

Qy      91  CPSCDSYEKKPPKPFLEFRKSLLOKMIHOHLSSRTHGSEDS 131
      || : : : : || : || : || : ||
Db      1075  C-----KESPRVLOPKFKLEEFFPKK--AKREHKMEVS 1106
      || : : : : || : || : || : ||

```

Search completed: May 23, 2001, 11:12:49
Job time: 252 sec

Result No.	Query #			DB	ID	Description
	Score	Match	Length			
1	101.5	14.6	162	1	IL15_FELCA	O97687 felis silve
2	99.5	14.3	162	1	IL15_HUMAN	P40933 homo sapien
3	94.5	13.6	162	1	IL15_CERAE	P40221 cercopitheci
4	94.5	13.6	162	1	IL15_MACMU	P48092 macaca mula
5	91.5	13.2	337	1	IRL1_MOUSE	P14719 mus musculu
6	90.5	13.0	162	1	IL15_BOVIN	Q28028 bos taurusu
7	82	11.8	607	1	ALB2_XENLA	P14872 xenopus lae
8	80.5	11.6	162	1	IL15_PIG	Q95253 sus scrofa
9	79	11.4	206	1	YIM7_YEAST	P04070 saccharomyc
10	78	11.2	741	1	IDH_AZOVI	P16100 azotobacter
11	74.5	10.7	441	1	FUS6_ARATH	P45432 arabidopsis
12	73.5	10.6	162	1	IL15_MOUSE	P48346 mus musculu
13	73.5	10.6	472	1	IFT2_HUMAN	P09913 homo sapien
14	73	10.5	304	1	IF2A_YEAST	P20459 saccharomyc
15	72.5	10.4	162	1	IL15_RAT	P97604 rattus norv
16	72.5	10.4	1206	1	FM14_MOUSE	Q05859 mus musculu
17	72	10.4	262	1	Y069_NPVAC	P41469 autographa
18	71.5	10.3	195	1	ALLA_YEAST	P32459 saccharomyc
19	71.5	10.3	484	1	CB3_MOUSE	P54285 mus musculu
20	71	10.2	334	1	LCMT_HUMAN	Q9u1c8 homo sapien
21	71	10.2	805	1	AHR_MOUSE	P30561 mus musculu
22	70.5	10.1	477	1	CB3_RABIT	P54286 oryctolagus
23	70.5	10.1	484	1	CB3_HUMAN	P54284 homo sapien
24	70.5	10.1	484	1	CB3_RAT	P54287 rattus norv
25	70	10.1	403	1	IFT3_MOUSE	Q64345 mus musculu
26	70	10.1	1222	1	YMP3_CAEEL	Q10947 caenorhabdi
27	69.5	10.0	864	1	YGM3_YEAST	P48237 saccharomyc
28	69.5	10.0	1468	1	FNW1_MOUSE	Q05860 mus musculu
29	68.5	9.9	464	1	N2B_HAIR	P46441 haematobia
30	68.5	9.9	2749	1	IF3R_CANFA	P29994 rattus norv
31	68	9.8	132	1	IL4_CANFA	O77762 canis famil
32	68	9.8	201	1	SDC2_HUMAN	P34741 homo sapien
33	67.5	9.7	300	1	SPY4_MOUSE	O9vtp2 mus musculu

```

RESULT 2
IL15_HUMAN
ID IL15_HUMAN STANDARD: PRT; 162 AA.
AC P40933; Q93058; O43512;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (IL15-S48AA ISOFORM).
RC TISSUE=Bone marrow;
RX MEDLINE=94233380; PubMed=8178155;
RA Grabstein K.K., Eisenman J., Sheanebeck K., Rauch C.,
RA Srinivasan S., Fung V., Beers C., Richardson J., Schoenborn M.A.,
RA Ahldeh M., Johnson L., Alderson M.R., Watson J.D., Anderson D.M.,
RA Giri J.G.;
RT "Cloning of a T cell growth factor that interacts with the beta chain
RT of the interleukin-2 receptor.";
RL Science 264:965-968(1994).
RN [2]
RP SEQUENCE FROM N.A. (IL15-S48AA ISOFORM).
RA Krause H., Jandrig B., Wernicke C., Bullone-Paus S., Pohl T.,
RA Diamantstein T.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (IL15-S21AA ISOFORM).
RC TISSUE=Lung cancer;
RX MEDLINE=96218668; PubMed=8668345;
RA Meazza R., Verdiani S., Blassoni R., Coppolecchia M., Gaggero A.,
RA Orenzo A.M., Colombo M.P., Azzone B., Ferrini S.;
RT "Identification of a novel interleukin-15 (IL-15) transcript isoform
RT generated by alternative splicing in human small cell lung cancer
RT cell lines.";
RL Oncogene 12:2187-2192(1996).
RN [4]
RP SEQUENCE FROM N.A. (IL15-S21AA ISOFORM).
RC TISSUE=Testis;
RX MEDLINE=98070771; PubMed=9405632;
RA Tagaya Y., Kurys G., Thies T.A., Losi J.M., Azimi N., Hanover J.A.,
RA Bamford R.N., Waldmann T.A.;
RT "Generation of secretable and nonsecretable interleukin 15 isoforms
RT through alternate usage of signal peptides.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:14444-14449(1997).
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA.
CC -!- SUBCELLULAR LOCATION: SECRETED (IL15-S48AA). IL15-S21AA IS NOT
CC SECRETED, BUT RATHER IS STORED INTRACELLULARLY, APPEARING IN THE
CC NUCLEUS AND CYTOPLASMIC COMPONENTS.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; IL15-S48AA (SHOWN HERE) AND
CC IL15-S21AA; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: MOST ABUNDANT IN PLACENTA AND SKELETAL MUSCLE.
CC IT IS ALSO DETECTED IN THE HEART, LUNG, LIVER AND KIDNEY. IL15-
CC S21AA IS PREFERENTIALLY EXPRESSED IN TISSUES SUCH AS TESTIS AND
CC THYMUS.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC -!- DATABASE: NAME=R&D Systems' cytokine source book;
CC WWW="http://www.rndsystems.com/cyt_cat/il-15.html".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U14407; AAA21551.1; -
DR EMBL; X91233; CAA62616.1; -
DR EMBL; X94223; CAA63914.1; -
DR EMBL; X94222; CAA63913.1; -
DR EMBL; AF031167; AAB97518.1; -
DR MIM; 600554; -
KW Cytokine; Glycoprotein; Signal; Alternative splicing.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48 POTENTIAL.
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1 37 MRISPHLRISITQCYLCVLCVLLNSHFRTEAGIHVFILG ->
FT MVLTGIDLC (IN ISOFORM IL15-S21AA).
FT CONFLICT 141 141 E -> K (IN REF. 4).
FT SEQUENCE 162 AA; 18086 MW; 0CE0520C1D8379E2 CRC64;
Query Match 14.3%; Score 99.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.0097;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;
QY 12 IDIVDQIKNYVNDLVPEF-----LPAPEDVETNCWSAFSCF----OKAOLKSANTGNE 62
DB 51 VNVISDLKK-IEDLIQSMHIDATLYTSDVHPSCVKYAMKCFLEQLQVLSLESDASIH 109
QY 63 RIINVSIIKKLKRKPPSTNAGRRKQHRLTCPSCDSYKPKPKFELRPSKLLQKMIH 118
DB 110 TVENLI--LANNLSNGNVTES---GCKECELEEKNIKEFLQSPFHVIVQMF 160
RESULT 3
ID IL15_CERAE STANDARD; PRT; 162 AA.
AC P40221;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 49-81.
RC TISSUE=Kidney;
RX MEDLINE=94233380; PubMed=8178155;
RA Grabstein K.H., Eisenman J., Sheanebeck K., Rauch C.,
RA Srinivasan S., Fung V., Beers C., Richardson J., Schoenborn M.A.,
RA Ahldeh M., Johnson L., Alderson M.R., Watson J.D., Anderson D.M.,
RA Giri J.G.;
RT "Cloning of a T cell growth factor that interacts with the beta chain
RT of the interleukin-2 receptor.";
RL Science 264:965-968(1994).
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC -----
DR EMBL: U03099; AAA18416.1;
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 127 127 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 162 AA; 18222 MW; 1BF9A82644E1C9B7 CRC64;

Query Match 13.6%; Score 94.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.029;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 12 IDIVDLKKNVNDLVPF-----LPAPEDVETNCWSAFSCF-----KQAQLKSANTGNNE 62
Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPKCVKVTAMKCFLELQVISHESGDTDIHD 109
QY 63 RIINVSIIKKLRKPPSTNAGROKRLTCPCSDSYEKPKPKFLEFKSLQKMIH 118
Db 110 TVENLII--LANNILSSNGNITES---GCKECEELEEKNIKEFLQSFVHIQVMEIN 160

RESULT 4
IL15_MACMU
ID IL15_MACMU STANDARD; PRT; 162 AA.
AC P48092;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=96003435; PubMed=7561102;
RA Villinger F.J., Brar S.S., Wayne A.E., Chikkala N., Ansari A.A.;
RT "Comparative sequence analysis of cytokine genes from human and
RL nonhuman primates.";
RL J. Immunol. 155:3946-3954(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Tatsumi M.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U19843; AAB60398.1;
DR EMBL; AB000555; BAA19149.1;
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48 POTENTIAL.
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133

```

```

FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 127 127 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARIANT 5 5 K -> T.
FT VARIANT 31 31 I -> T.
SQ SEQUENCE 162 AA; 18194 MW; D233CF7FF6188C01 CRC64;

Query Match 13.6%; Score 94.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.029;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 12 IDIVDLKKNVNDLVPF-----LPAPEDVETNCWSAFSCF-----KQAQLKSANTGNNE 62
Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPKCVKVTAMKCFLELQVISHESGDTDIHD 109
QY 63 RIINVSIIKKLRKPPSTNAGROKRLTCPCSDSYEKPKPKFLEFKSLQKMIH 118
Db 110 TVENLII--LANNILSSNGNITES---GCKECEELEEKNIKEFLQSFVHIQVMEIN 160

RESULT 5
IRL1_MOUSE
ID IRL1_MOUSE STANDARD; PRT; 337 AA.
AC P14719;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE INTERLEUKIN 1 RECEPTOR-LIKE 1 PRECURSOR (ST2 PROTEIN) (T1 PROTEIN)
DE (LYMPHOCYTE ANTIGEN 84).
GN IL1RL1 OR ST2 OR ST2 OR LY84.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=90092495; PubMed=2532153;
RA Tomimaga S.;
RT "A putative protein of a growth specific cDNA from BALB/c-3T3 cells
RT is highly similar to the extracellular portion of mouse interleukin 1
RT receptor.";
RN FEBS Lett. 258:301-304(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/HE; TISSUE=Spleen;
RX MEDLINE=91355215; PubMed=1832015;
RA Tomimaga S.I., Jenkins N.A., Gilbert D.J., Copeland N.G.,
RA Tetsuka T.;
RT "Molecular cloning of the murine ST2 gene. Characterization and
RT chromosomal mapping.";
RL Biochim. Biophys. Acta 1090:1-8(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89345536; PubMed=2527364;
RA Klemenz R., Hoffmann S., Werenskiold A.K.;
RT "Serum- and oncoprotein-mediated induction of a gene with sequence
RT similarity to the gene encoding carcinoembryonic antigen.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:5708-5712(1989).
CC -!- FUNCTION: POSSIBLY INVOLVED IN REGULATION OF T-LYMPHOCYTE
CC ACTION.
CC -!- DEVELOPMENTAL STAGE: GROWTH-SPECIFIC EXPRESSION, G1-PHASE OF
CC CELL CYCLE.
CC -!- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -!- SIMILARITY: STRONG. TO INTERLEUKIN-1 RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL; Y07519; CAA68812.1; -
DR DR EMBL; X60184; CAA42742.1; -
DR EMBL; M24843; AAA40160.1; -
DR PIR; S07054; S07054.
DR MGD; MGI:98427; L984.
DR InterPro: IPR003006; -
DR Pfam; PF00047; Ig; 3.
KW Immunoglobulin domain; Glycoprotein; Signal.
FT SIGNAL 1 26
FT CHAIN 27 337 INTERLEUKIN 1 RECEPTOR-LIKE 1.
FT DOMAIN 35 100 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 132 194 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 233 315 IG-LIKE C2-TYPE DOMAIN.
FT FT DISULFID 42 93 BY SIMILARITY.
FT FT DISULFID 139 187 BY SIMILARITY.
FT FT DISULFID 240 308 BY SIMILARITY.
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 107 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 225 225 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 192 192 A -> V (IN STRAIN C3H/HE).
FT SEQUENCE 337 AA; 38502 MW; 7574372722486926 CRC64;
Query Match 13.2%; Score 91.5; DB 1; Length 337;
Best Local Similarity 25.6%; Pred. No. 0.13;
Matches 33; Conservative 20; Mismatches 39; Indels 37; Gaps 7;
QY 16 DLKKNYVNDLPPELPAPEDVETNCWSAFSCFOKAQLKSANTGNERNINVSIKLKRK 75
DB 76 DRUK-----FLPARVE-----DSGIVACV-----IRSPNL-NK7GYLNVTTIHK---K 113
QY 76 PPSTN-----ACGRQKRLHRTCPSCDSYERKKPKPELRFKSLQKMIHQH----- 120
DB 114 PPSGNIPDYLMSYTVRSGDNFKITCTPTIDLYNWTAPVQMFKNCKALQEPFRAHSYLF 173
QY 121 LSRTHGSE 129
DB 174 IDNVTHDE 182
RESULT 6
IL15_BOVIN
ID IL15_BOVIN STANDARD; PRT; 162 AA.
AC Q28028;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RS SEQUENCE FROM N.A.
ST STRAIN=HOLSTEIN;
RX MEDLINE=97426124; PubMed=9282828;
RA Canals A., Gasbarre L.C., Boyd P.C., Almeria S., Zarlenga D.S.;
RT *Cloning and expression of bovine interleukin-15: analysis and
RT modulation of transcription by exogenous stimulation.*;
RL J. Interferon Cytokine Res. 17:473-480(1997).
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES THE INTERACTION OF T-
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-15
CC GAMMA BUT NOT IL-2R ALPHA (BY SIMILARITY).

```


Qy	94	CDSEYKPKPKEFLERFKSLLOKQMHOHLSRTHGSED	130
	:	: : :	:
	:	: : :	:
Dd	192	CEEDKE--KCFSEKMQLMK-----QSHSID	217

		RESULT	8
ID	IL15_PIG		
ID	IL15_PIG	STANDARD;	PRT; 162 AA.
AC	Q95253;		
DT	01-NOV-1997	(Rel. 35, Created)	
DT	01-NOV-1997	(Rel. 35, Last sequence update)	
DT	15-DEC-1998	(Rel. 37, Last annotation update)	
DE	INTERLEUKIN-15	PRECURSOR (IL-15).	
GN	IL15.		
OS	Sus scrofa	(Pig).	
OC	Eukaryota;	Metazoa;	Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia;	Eutheria;	Cetartiodactyla; Suina; Suidae; Sus.
OX	NCBI_TaxID=9823;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Blood;		
RX	MEDLINE=97449311;	PubMed=9305780;	
RA	Canals A., Grimm D.R., Gasbarre L.C., Lunney J.K., Zarlenga D.S.;		
RT	"Molecular cloning of cDNA encoding porcine interleukin-15.";		
RL	Gene 195;337-339(1997).		
CC	-I- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15 WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R GAMMA BUT NOT IL-2R ALPHA (BY SIMILARITY).		
CC	-II- SUBCELLULAR LOCATION: SECRETED.		
CC	-III- SIMILARITY: BELONGS TO THE IL-15 FAMILY.		

```

Query Match          11.6%; Score 80.5; DB 1; Length 162;
Best Local Similarity 24.2%; Pred. No. 0.6;
Matches 29; Conservative 20; Mismatches 44; Indels 27; Gaps 6;

QY   14  IVDDKLVVDLVEPF-----LPAPEDVETNCNSAFSCF-----QRAQLKSANTG 59
      :: || : ||: | | : ||: | ||: | ||: | | : ||: | |
Db    53  VISDLKK-IEDLIIRSIHMDATLYTESDAHPNCKYTKAMCFLELRVLQESRNSDISDTV 111
      :: ||: | |:: | |:: | |:: | |:: | |:: | |:: | |:: | |:: | |:: | |
QY   60  NNERII-NVSTIKLKRRPPSTNAGRRQKHRLTCPCSDSYEKKKPKPEFLERKSLLOKMIH 118
      -| -| : |:: | |:: | |:: | |:: | |:: | |:: | |:: | |:: | |:: | |
Db    112 ENLLILANSLSISIEYK---TESG-----CAECELEEKNFEFLKSIHIHQVFIN 160

RESULT 9
YIM7 YEAST
ID YIM7 YEAST STANDARD; PRT; 206 AA..
AC P40470;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 23.8 KDA PROTEIN IN MET18-STHL INTERGENIC REGION.
GN YIM127C.
```


Matches	22;	Conservative	30;	Mismatches	54;	Indels	5;	Gaps	3;
QY	12	IDI----	VDLKNVYNDL-VPEFLPAPEDVETNCEWSAFSCFOAKLSKANTGNNERIINV	67					
Db	51	IDVRYDEKTESLIQSHIDTTLTYDSDFFHPSCVKVTAMNCF-LLELOQVILHEYSNM	109						
QY	68	SIKRKLKPPSTNAGRROKHLRTCPSCDSYKPKPPKFLERKSLSLQKMIH	118						
Db	110	TVRNVLANSSTLSSNKNVAESGCKECEELEKXFTTEFLQSFIRIVQMFIN	160						
RESULT	13								
IFT2_HUMAN									
ID	IFT2_HUMAN	STANDARD;	PRT;	472	AA.				
AC	P09113;								
DT	01-MAR-1989	(Rel. 10, Created)							
DT	01-MAR-1989	(Rel. 10, Last sequence update)							
DT	01-OCT-2000	(Rel. 40, Last annotation update)							
DE	INTERFERON-INDUCED PROTEIN WITH TETRAVIRICOPEPTIDE REPEATS 2 (IFIIT-2)								
DE	INTERFERON-INDUCED 54 KDA PROTEIN (IFI-54K) (ISG-54 K).								
GN	IFIIT2 OR IFI54 OR G10P2.								
OS	Homo sapiens (Human).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
NCBI_TaxID	9606;								
RP	[1]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=87067427; PubMed=3466167;								
RA	Levy D., Larner A., Chaudhuri A., Babis L.E., Darnell J.E. Jr.;								
RT	"Interferon-stimulated transcription: isolation of an inducible gene								
RT	and identification of its regulatory region."								
RL	Proc. Natl. Acad. Sci. U.S.A. 83:8929-8933(1986).								
RN	[2]								
RP	SEQUENCE OF 1-2 FROM N.A.								
RX	MEDLINE=88254799; PubMed=2454816;								
RA	Wathelet M.G., Claus I.M., Content J., Huez G.A.;								
RT	"Regulation of two interferon-inducible human genes by interferon,								
RT	poly(rI).poly(rC) and viruses."								
RL	Eur. J. Biochem. 174:323-329(1988).								
RN	[3]								
RP	SIMILARITY TO IFI-56K.								
RX	MEDLINE=88196376; PubMed=3360121;								
RA	Wathelet M.G., Claus I.M., Content J., Huez G.A.;								
RT	"The IFI-56K and IFI-54K interferon-inducible human genes belong to								
RT	the same gene family."								
RL	FEBS Lett. 231:164-171(1988).								
CC	-I- INDUCTION: BY INTERFERONS.								
CC	-I- SIMILARITY: BELONGS TO THE IFIT FAMILY.								
CC	-I- SIMILARITY: CONTAINS TPR REPEATS.								
CC									
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/								
CC	or send an email to license@isb-sib.ch).								
CC									
EMBL; M14660; AAA59191.1;									
EMBL; M14659; AAA59191.1; JOINED.									
EMBL; X07557; CAA30438.1;									
DR	MIM; 147040;								
KW	Repeat; TPR repeat; Interferon induction.								
FT	REPEAT	51	TPR 1.						
FT	REPEAT	87	TPR 2.						
FT	REPEAT	157	TPR 3B.						
FT	REPEAT	175	TPR 4.						
FT	REPEAT	210	TPR 5.						
FT	REPEAT	248	TPR 6.						

FT PROPEP	30	48	POTENTIAL.
FT CHAIN	49	162	INTERLEUKIN-15.
FT DISULFID	83	133	POTENTIAL.
FT DISULFID	90	136	POTENTIAL.
FT CARBOHYD	10	10	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	104	104	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	108	108	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	119	119	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	143	143	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC	81	96	MISSING (IN SHORT ISOFORM).
SQ SEQUENCE	162 AA;	18804 MW;	DA68097A83065AEA CRC64;

Query Match 10.4%; Score 72.5; DB 1; Length 162;
Best Local Similarity 18.8%; Pred. No. 3.4;
Matches 22; Conservative 32; Mismatches 60; Indels 3; Gaps 2;

QY 2 DRHMIRMROLIDIVDOLKNYVNDLVPEFLPAPEDVETNCWSEAFSCFQKAQLKASANTGNN 61
Db 47 EANNIDVRYDLEKIESLIQFIH--IDTTLTSDDFHPSCVKVTAMNCF-LLELQVILHEYS 103

QY 62 ERIINVSIRKKLRKPPSTNAGRQKHRLTCPSCDSYEKKPPKEFLERFKSLLOKMIH 118
Db 104 NMTLNETVRNVLYLANSTLSSNKNVIESGCKECEELEERNFTFELOSFIHVOMFIN 160

Search completed: May 23, 2001, 11:20:24
Job time: 522 sec

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	99.5	14.3	114	4	Q0UBA3		Q0UBA3 homo sapien
2	99.5	14.3	136	4	Q00440		Q00440 homo sapien
3	91.5	13.2	567	11	Q05208		Q05208 mus musculus
4	89.5	12.9	162	6	Q9XSJ6		Q9XSJ6 ovine aries
5	83.5	12.0	3052	14	Q82933		Q82933 johnsongras
6	81.5	11.7	566	5	Q9VYE8		Q9VYE8 drosophila
7	75.5	10.9	336	11	Q62612		Q62612 rattus norv
8	75.5	10.9	566	11	Q62611		Q62611 rattus norv
9	74.5	10.7	424	5	Q23239		Q23239 calliphora
10	74	10.6	187	13	Q9W756		Q9W756 gallus gall
11	74	10.6	334	4	Q9Y378		Q9Y378 homo sapien
12	74	10.6	440	10	Q9M2E9		Q9M2E9 arabidopsis
13	74	10.6	2197	5	Q96296		Q96296 plasmodium
14	73.5	10.6	868	5	Q9NAH8		Q9NAH8 caenorhabdi
15	73	10.5	848	4	Q75141		Q75141 homo sapien
16	72.5	10.4	305	11	Q9J048		Q9J048 mus musculus
17	72.5	10.4	309	14	Q92504		Q92504 bombyx mori
18	72.5	10.4	1101	5	Q62471		Q62471 caenorhabdi
19	72	10.4	257	1	Q27054		Q27054 methanobact

DT 01-JUL-1997 (TEMBLrel. 04, Last sequence update)
DT 01-AUG-1998 (TEMBLrel. 07, Last annotation update)
DE INTERLEUKIN-15.
GN IL-15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Meazza R., Ferrini S.;
RL Submitted (APR-1997) to the EMBL/GenBank/DDJB databases.
DR EMBL; Y09908; CAA1044.1; -;
SQ SEQUENCE 136 AA; 15259 MW; 494ED8DDC5D364CD CRC64;

```

Query Match          14.3%; Score 99.5; DB 4; Length 136;
Best Local Similarity 25.9%; Pred. No. 0.01;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

```

QY 12 IDVDLNKYNVNDLPVEF-----LPAPDEVETNCWEAFSCF----OKAOLKSANTGNE 62
DB 25 VNVISDLKK-IEDIQSMDHIDAITYTSDVHPVCCKVTAMKFLLQLVLVISLEGDSASHD 83
QY 63 RIINVSIKKLRPPSTNAGROKHRTCTPCSDSYEKPKPEFLERFKSLQKMIH 118
DB 84 TVENLI--LANSSLSNGNWTES---GCKECEBLEEKNKEFLQSFVHVQMFIN 134

RESULT 3
Q05208 PRELIMINARY; PRT; 567 AA.

ID.	Q05208
AC	Q05208;
DT	01-JUN-1998 (TEMBLrel. 06, Created)
DF	01-JUN-1998 (TEMBLrel. 06, Last sequence update)
DT	01-JUN-2000 (TEMBLrel. 14, Last annotation update)
DE	ST2L PROTEIN PRECURSOR.
GN	LY84.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	MEDLINE=93170492; PubMed=7916701;
RX	Yanaqisawa K., Takagi T., Tsukamoto T., Tetsuka T., Tomimaga S.;
RT	"Presence of a novel primary response gene ST2L, encoding a product highly similar to the interleukin 1 receptor type 1.";
RL	FEBS Lett. 318:83-87(1993).
DR	EMBL; D13695; BAA02854.1; -;
DR	MGD; MG1:98427; LY84.
DR	INTERPRO; IPR000157; -;
DR	INTERPRO; IPR002052; -;
DR	INTERPRO; IPR003006; -;
DR	Pfam; PF00047; Ig; 3.
DR	Pfam; PF01582; TIR; 1.
DR	PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW	SIGNAL.
FT	SIGNAL.
FT CHAIN	1 26 POTENTIAL.
FT SIGNAL	27 567 ST2L PROTEIN.
SQ	SEQUENCE 567 AA; 64801 MW; 9228CE227D95B0BC CRC64;

Query Match 13.2%; Score 91.5; DB 11; Length 567;
Best Local Similarity 25.6%; Pred. No. 0.3;
Matches 33; Conservative 20; Mismatches 39; Indels 37; Gaps 7;

QY 16 DLKNKYNDLVPEFLPAPDEVETNCWEAFSCFOKAOLKSANTGNNRIINVSIKKLRK 75
DB 76 DRLK-----FLPARVE-----DSGIYACV----IRSPL-NKTCYLNVTHK----K 113
QY 76 PPSTN-----AGGRQKHRTCTPCSDSYEKPKPEFLERFKSLQKMIHQH----- 120
DB 76 DRLK-----AGGRQKHRTCTPCSDSYEKPKPEFLERFKSLQKMIHQH----- 120

DR MEROPS: S30.001; -
 DR INTERPRO: IPR001205; -
 DR INTERPRO: IPR001410; -
 DR INTERPRO: IPR001456; -
 DR INTERPRO: IPR001592; -
 DR INTERPRO: IPR001730; -
 DR INTERPRO: IPR002540; -
 DR PFAM: PF00270; DEAD; 1.
 DR PFAM: PF00680; RNA_dep_RNA_pol; 1.
 DR PFAM: PF00767; Pory_coat; 1.
 DR PFAM: PF00851; Peptidase_C6; 1.
 DR PFAM: PF00863; Peptidase_C4; 1.
 DR PFAM: PF01577; Pory_PL; 1.
 DR PRINTS: PR00966; NIAPOTYPTASE.
 KW Coat protein; Protease.
 FT CHAIN 239 699
 FT CHAIN 700 1046
 FT CHAIN 1047 1098
 FT CHAIN 1099 1757
 FT CHAIN 1758 1803
 FT CHAIN 1804 1991
 FT CHAIN 1992 2233
 FT CHAIN 2234 2749
 FT CHAIN 2750 3052
 SQ SEQUENCE 3052 AA; 347249 MW; 08CD8831A73EBCA9 CRC64;

Query Match 12.0%; Score 83.5; DB 14; Length 3052;
 Best Local Similarity 23.1%; Pred. No. 12;
 Matches 25; Conservative 23; Mismatches 57; Indels 3; Gaps 3;
 QY 14 IVDLKKNVNDLPPEFLPAPDEVTNCEWSAFCFQKAQLKSAFTGNNGNININVSIKLK 73
 Db 220 LVNALDQYED-VKQICHYFSDAERAFWKGFTENHTAQREADHNNHPV-MSVECG 277
 QY 74 RKPESTWAGRRQKRLTSPC-DSYEKKPKPEFLERFKSLQKMIHQH 120
 Db 278 RRAAMLENAFHQGFKITCKHCFTQFDEHSDEVCERIHNALQRIEEQN 325

RESULT 6
 Q9VYB8 PRELIMINARY; PRT; 566 AA.
 AC Q9VYB8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE CG15745 PROTEIN.
 GN CG15745.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=107311132;
 RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananaitides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballev R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottilier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadenou E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahike S., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of Drosophila melanogaster";
 FT Science 287:2183-2195(2000).
 RL Science 287:2183-2195(2000).
 DR EMBL; AE003492; AAF48250.1; -
 DR FLYBASE; FBgn0030469; CG15745.
 SQ SEQUENCE 566 AA; 60620 MW; 31FA41FA4A4EF427F CRC64;

Query Match 11.7%; Score 81.5; DB 5; Length 566;
 Best Local Similarity 26.1%; Pred. No. 3.1;
 Matches 24; Conservative 21; Mismatches 40; Indels 7; Gaps 3;
 QY 35 DVEINCEWSAFCFQKAQLKSAFTGNNGNININVSIIKKRPSTNAGRQKRLTSP 92
 Db 120 DIGPCEVASISELALRKAQLKAQFFGNOVG---GLARDSETSTRITRTTNTYRSAYP 175
 QY 93 SCDSYEEKPKPEFLERFKSLQKMIHQHLS 123
 Db 176 SKTERGFPVQLIDQFQAMIVQQQQQLSN 207

RESULT 7
 Q62612 PRELIMINARY; PRT; 336 AA.
 ID Q62612
 AC Q62612;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE FIT-1S PRECURSOR.
 GN FIT-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94178260; PubMed=8131748;
 RA Bergers G., Reiksterfer A., Braselmann S., Graninger P.,
 RA Busslinger M.;
 RT "Alternative promoter usage of the Fos-responsive gene Fit-1 generates
 RT mRNA isoforms coding for either secreted or membrane-bound proteins
 RT related to the IL-1 receptor";
 RL EMBO J. 13:1176-1188(1994).
 DR EMBL; U04319; AAA67172.1; -
 DR INTERPRO; IPR003006; -
 DR PFAM; PF00047; Ig; 3.
 KW Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 336
 SQ SEQUENCE 336 AA; 38090 MW; 61C6A6FA8FE319D2 CRC64; FIT-1S.

Query Match 10.9%; Score 75.5; DB 11; Length 336;
Best Local Similarity 22.2%; Pred. No. 7.1; Mismatches 26; Indels 37; Gaps 6;
Matches 30; Conservative 26; Mismatches 26; Indels 37; Gaps 6;

QY 12 IDIVDLKKNVNDLVP-----EFLPAPEDVETNCWSAFSCFQKALKSAN 57
DB 49 INPVEWYISNTNERIPTQKRNRIFVSRDLKFLPAKVE-----DSGIYTCVIRSP-ESIK 102

QY 58 TGNNERIINVSIKLKRPP-----STNAGRRQKRLTFCSDSYEKPKPFLE 107
DB 103 TGS-----LNVITYK---RPPNCKIPDYMYSTVDGSDKNSKITCPTIALYNTAPVQWFK 155

QY 108 RPKSLQKMIHOHLS 122
DB 156 NCKALQGFRAHMS 170

RESULT 8
Q62611 PRELIMINARY; PRT; 566 AA.
AC Q62611;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE FIT-1M PRECURSOR.
GN FIT-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94178260; PubMed=8131748;
RA Bergers G., Reiksterfor A., Braselmann S., Graninger P.,
RA Busslinger M.;
RT "Alternative promoter usage of the Fos-responsive gene Fit-1 generates
RT mRNA isoforms coding for either secreted or membrane-bound proteins
RT related to the IL-1 receptor."
RL EMBO J. 13:1176-1188(1994).
DR EMBL; U04317; AAA18480.1;
DR INTERPRO; IPR000157;
DR PFAM; PF00047; 1g; 3;
DR PFAM; PF01582; TIR; 1.
KW Signal.
FT SIGNAL 1 26
FT CHAIN 27 566
SQ SEQUENCE 566 AA; 64405 MW; 50554A29437A0C31 CRC64;

Query Match 10.9%; Score 75.5; DB 11; Length 566;
Best Local Similarity 22.2%; Pred. No. 12;
Matches 30; Conservative 26; Mismatches 26; Indels 37; Gaps 6;

QY 12 IDIVDLKKNVNDLVP-----EFLPAPEDVETNCWSAFSCFQKALKSAN 57
DB 49 INPVEWYISNTNERIPTQKRNRIFVSRDLKFLPAKVE-----DSGIYTCVIRSP-ESIK 102

QY 58 TGNNERIINVSIKLKRPP-----STNAGRRQKRLTFCSDSYEKPKPFLE 107
DB 103 TGS-----LNVITYK---RPPNCKIPDYMYSTVDGSDKNSKITCPTIALYNTAPVQWFK 155

QY 108 RPKSLQKMIHOHLS 122
DB 156 NCKALQGFRAHMS 170

RESULT 9
Q23239 PRELIMINARY; PRT; 424 AA.
AC Q23239;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE YOLK PROTEIN 3 PRECURSOR.
OS Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestroidea; Calliphoridae; Calliphora.
OX NCBI_TaxID=7373;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94275878; PubMed=8007002;
RA Martinez A., Bowles M.;
RT "The sequence and expression pattern of the Calliphora erythrocephala
RT yolk protein A and B genes."
RL J. Mol. Evol. 38:336-351(1994).
DR EMBL; X70795; CAA50066.1;
DR INTERPRO; IPR000734;
DR PFAM; PF00151; Lipase; 1.
KW Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 424 YOLK PROTEIN 3.
SQ SEQUENCE 424 AA; 46439 MW; C35ECB971311603ED CRC64;

Query Match 10.7%; Score 74.5; DB 5; Length 424;
Best Local Similarity 25.7%; Pred. No. 11;
Matches 35; Conservative 20; Mismatches 48; Indels 33; Gaps 7;

QY 8 MRQLIDIVDLKKNVNDLVPFELPAPED-----VETNCWSAFSCFQKALKS 55
DB 66 MQKLDLHLSQIN---NDLAFSLPSSNVPCYIVKPKVSTSLDGLASACKQO----- 117

QY 56 ANTGNNERIINVSIKLKRPPSTNAGRRQKRLTFCSDSYEKPKPFLEFKSLQK 115
DB 118 PNFGEEL--VTILITGL---PATETVRKANRLIDAYLQRYSTK--RQPSKFDYSSEK 170

QY 116 MIHQHLSRTHSEDS 131
DB 171 M-----ARTSEEDS 180

RESULT 10
Q9W756 PRELIMINARY; PRT; 187 AA.
AC Q9W756;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (INTERLEUKIN 15 PRECURSOR).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCW; TISSUE=LIVER;
RA Burnside J., Sofer L.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SC; TISSUE=LIVER;
RA Choi K.D., Lillehoj H.S., Burnside J.;
RT "Gallus gallus mRNA for IL-15 precursor."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152927; RAD38392.1;
DR EMBL; AF139097; AAF61446.1;
SQ SEQUENCE 187 AA; 21964 MW; 470601EBF8837095 CRC64;

Query Match 10.6%; Score 74; DB 13; Length 187;
Best Local Similarity 22.5%; Pred. No. 5.4;
Matches 29; Conservative 16; Mismatches 46; Indels 38; Gaps 5;

QY 1 QDRMIRMQLIDIVDLQKNVNDLVPEFLPAPEDVETNCWSAFSCF---QKAQLKSAN 57
 Db 76 KDLKLEKTSIDIVSLYTAN-----TIEDIE--COEPMRCFFLEMKVILHECD 122
 QY 58 TGNRRIINVSIIKKLRKPPSTNAGRQRHRLT-----CPSCDSYEKKPKPKFL 106
 Db 123 -----IKKCRKHVDNRWNGNARFATYQLNSTAKKCKECEYEKKNFTFI 171
 QY 107 ERPKSLLOK 115
 Db 172 QSEVKVIOR 180

RESULT 11
 QY378 ID QY378 PRELIMINARY; PRT; 334 AA.
 AC QY378;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
 DE CGI-68 PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lin W.-C.;
 RT "Comparative gene cloning: Identification of novel human genes with
 RL Caenorhabditis elegans proteome as template.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF151826; AAD34063.1;
 SQ SEQUENCE 334 AA; 38279 MW; 3AD02031AC6C83DE CRC64;

Query Match 10.6%; Score 74; DB 4; Length 334;
 Best Local Similarity 26.4%; Pred. No. 10;
 Matches 42; Conservative 22; Mismatches 55; Indels 40; Gaps 10;

QY 8 MRQLIDIVDLQKN-VYNDLVPEFLP-----PEDVETNCWSAFSCFQKAQL---KS 55
 Db 172 LRDLSELEKLRKCNNTQPLLLIAECVLVYMPQSANLLKWAANS-FERAMFNIYEQ 230

QY 56 ANTGNRRIINVSIIKKLRKPPSTNAG-----RRQKHL-----TCPSCDSYE--KK 100
 Db 231 VNMG--DREGQIMIEHL-RRPSCDLAGVETCKSLQSKERLLSNGWETASAVDMMLYNR 287

QY 101 PKP-----EFLERFKSLLOKMIHQHLSRTHGSED 130
 Db 288 LPRAEVSRIESLEFLDEMLLEQLMRHYCLCWATKGGNE 326

RESULT 12
 QY2E9 ID QY2E9 PRELIMINARY; PRT; 440 AA.
 AC QY2E9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE FUSCA PROTEIN FUS6.
 GN T20K12.40.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA De Haan M., Maarse A.C., Grivell L.A., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X., Quetier F., Emlenoubat M.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL137898; CAB71044.1;
 SQ SEQUENCE 440 AA; 50441 MW; 4598E9D2485C586A CRC64;

Query Match 10.6%; Score 74; DB 10; Length 440;
 Best Local Similarity 22.0%; Pred. No. 13;
 Matches 26; Conservative 22; Mismatches 50; Indels 20; Gaps 3;

QY 3 RHMIRMQLIDIVDLQKNVNDLVPEFLPAPEDVETNCWSAFSCFQKAQLKSANTGNNE 62
 Db 230 KYKLAARKFLDVNPELNGSYNEVI-----APQDIATYGGLCALASDFRSEKAF 278

QY 63 RIINVSIIKKLRKPPSTNAGRQRHRLTCTPCSDSYEKKPKPKFLERFKSLLOKMIHQH 120
 Db 279 -IDNINFRNFLELVPDVRELINDFYSSRYASC-----LEYLASKSNLLLDIHLH 327

RESULT 13
 OY6296 ID OY6296 PRELIMINARY; PRT; 2197 AA.
 AC OY6296;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
 DE PFEML.
 GN PFB1055C.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99021743; PubMed=9804551;
 RA Gardner M.J., Tettelin H., Carucci D.J., Cummings L.M., Aravind L.,
 RA Koonin E.V., Shallow S., Mason T., Yu K., Fujii C., Pederson J.,
 RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M.,
 RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
 RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
 RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
 falciparum.";
 RT Science 282:1126-1132(1998).
 RL Science 282:1126-1132(1998).
 DR EMBL; AE001434; AAC71996.1;
 SQ SEQUENCE 2197 AA; 249668 MW; D2A92DB54535C143 CRC64;

Query Match 10.6%; Score 74; DB 5; Length 2197;
 Best Local Similarity 29.5%; Pred. No. 74;
 Matches 28; Conservative 16; Mismatches 41; Indels 10; Gaps 4;

QY 18 LKNVNDLVPEFLPAPEDVETNCWSAF--SCFQKAQLKSANTGNRRIINVSIIKKLRK 75
 Db 1594 VKSPLETWIPK-IAVNDQDNVILKSPGNSCGCSAISTN-GNEIDAIDCMIKKLEKK 1651

QY 76 -----PPSTNAGRQRHRLTCTPCSDSYEKKPKPE 104
 Db 1652 IDECKRKPGENSGQTCTNCTLTHPLDVOQDEPLEE 1686

RESULT 14
 QYNAH8 ID QYNAH8 PRELIMINARY; PRT; 868 AA.
 AC QYNAH8;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE Y47D3A.14 PROTEIN.
 GN Y47D3A.14.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]

```

RP SEQUENCE FROM N.A.
RA Matthews L.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AL117202; CAB55073.1; -.
SQ SEQUENCE 868 AA; 100165 MW; 09A2A97066E373D7 CRC64;

Query Match 10.68; Score 73.5; DB 5; Length 868;
Best Local Similarity 29.2%; Pred. No. 31;
Matches 19; Conservative 11; Mismatches 30; Indels 5; Gaps 1;

Qy 42 WSASFQKAOGLKS-----ANTGNNERIINVSIIKKLRKPPSTNAGRRKRLTCPCSDS 96
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 680 WSRVQAFKKKMKKSGGGTSSNSDSESEKVKLKKARPSDETIVIPVPHYIVCPLPFS 739
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 97 YEKKP 101
   : |
Db 740 HVAIP 744

RESULT 15
O75141
ID O75141 PRELIMINARY; PRT; 848 AA.
AC O75141;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE KIAA0650 PROTEIN (FRAGMENT).
GN KIAA0650.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=98403880; PubMed=9734811;
RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. X.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 5:169-176(1998).
DR EMBL; AB014550; BAA31625.1; -.
FT NON_TER 1
SQ SEQUENCE 848 AA; 95756 MW; 2192522F40ED9E7B CRC64;

Query Match 10.58; Score 73; DB 4; Length 848;
Best Local Similarity 25.5%; Pred. No. 34;
Matches 27; Conservative 22; Mismatches 47; Indels 10; Gaps 5;

Qy 28 EFLPAPEDVETNC-EWSAFSCFQKAOGLKSANTGNNERIINVSIIKKLRKPPSTNAGRRK 86
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 66 KFIQPPGNKDLCTWRFEFSDFIRVOLIS---GPPAKLLLDWPELKESIPVIN-GRDLQ 121
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 87 HRLTCPCSDSYEKPKPEFLERFKSLQAKMIHQHL--SSRTHGSED 130
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 122 NPIIVQLCDQNDNPAP---VQHVKISLTAKSNLKLMPNSQQHKTDE 164
   : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: May 23, 2001, 11:19:41
Job time: 594 sec

GenCore version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:11:34 ; Search time 108.07 Seconds
(without alignments)
55.539 Million cell updates/sec

Title: US-09-522-217-2_COPY_41_145

Perfect score: 557

Sequence: 1 QLIDIVDLKYNVNDLPDF.....DSYEKKPKPEFLERFKSLIQ 105

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0401.*

1:	/SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2:	/SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3:	/SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4:	/SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5:	/SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6:	/SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7:	/SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8:	/SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9:	/SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10:	/SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11:	/SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12:	/SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13:	/SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14:	/SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15:	/SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16:	/SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17:	/SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18:	/SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19:	/SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20:	/SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21:	/SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22:	/SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	557	100.0	162 21 B18623	A human zalphall 1
2	557	100.0	519 21 B18627	Amino acid sequenc
3	325	58.3	146 21 B18624	A mouse zalphall 1
4	325	58.3	510 21 B18628	Amino acid sequenc
5	164	29.4	40 21 B18625	Antigeninc peptide
6	133	23.9	32 21 B18626	Antigeninc peptide
7	99.5	17.9	135 21 Y54825	Human Interleukin-
8	95.5	17.1	114 16 R83310	Human Interleukin-
9	95.5	17.1	114 16 R83435	Mammalian interleu
10	95.5	17.1	114 16 R66928	Mammalian IL-15.
11	95.5	17.1	114 17 W09099	Simian mature epit

12	95.5	17.1	114	17	W07253	Generic mammalian
13	95.5	17.1	114	19	W39186	Simian epithelium
14	95.5	17.1	114	20	W39187	Simian epithelium
15	95.5	17.1	114	21	Y52309	Mature simian epit
16	95.5	17.1	162	16	R83438	Human interleukin-
17	95.5	17.1	162	16	R66927	Human IL-15. Homo
18	95.5	17.1	162	17	W09098	Simian epithelium
19	95.5	17.1	162	17	W07255	Human epithelium-d
20	95.5	17.1	162	18	R98527	Human interleukin-
21	95.5	17.1	162	18	W37369	Wild-type interleu
22	95.5	17.1	162	18	W01658	Human interleukin-
23	95.5	17.1	162	19	W53878	Human interleukin-
24	95.5	17.1	162	19	W39185	Simian epithelium
25	95.5	17.1	162	20	Y03756	Simian epithelium-
26	95.5	17.1	162	21	B18632	Amino acid sequenc
27	95.5	17.1	162	21	W78595	Human interleukin-
28	95.5	17.1	162	21	Y52308	Simian epithelium-
29	95.5	17.1	162	22	B50870	Human IL-15. Homo
30	90.5	16.2	114	16	R83309	Simian interleukin
31	90.5	16.2	114	17	W09101	Human mature epit
32	90.5	16.2	114	19	W39188	Human epithelium d
33	90.5	16.2	114	20	Y03759	Human epithelium-d
34	90.5	16.2	114	21	Y52311	Mature human epit
35	90.5	16.2	122	17	R90842	Recombinant flag s
36	90.5	16.2	162	16	R83436	Simian interleukin
37	90.5	16.2	162	16	R66926	Simian IL-15. Cer
38	90.5	16.2	162	17	W09100	Human epithelium d
39	90.5	16.2	162	17	W07254	Simian epithelium-
40	90.5	16.2	162	17	R98526	Simian interleukin
41	90.5	16.2	162	17	R92798	Mammalian interleu
42	90.5	16.2	162	18	W37370	Mutant interleukin
43	90.5	16.2	162	19	W39187	Human epithelium d
44	90.5	16.2	162	20	Y03758	Human epithelium-d
45	90.5	16.2	162	21	Y78594	Simian interleukin

ALIGNMENTS

RESULT 1

ID B18623 standard; Protein; 162 AA.

AC B18623;

DT 22-JAN-2001 (first entry)

XX A human zalphall ligand polypeptide.

DE zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
tumourigenesis; leukaemia; hematopoiesis; B cell tumour.

XX Homo sapiens.

OS WO2000053761-A2.

PN 14-SEP-2000.

PD 09-MAR-2000; 2000WO-US06067.

PF 09-MAR-1999; 99US-0264908.

PR 11-MAR-1999; 99US-0265992.

PR 01-JUL-1999; 99US-0142013.

XX (ZYMO) ZYMOGENETICS INC.

PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Gross JA;

PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;

XX WPI; 2000-565600/52.

DR N-PSDB; A75552.

XX New human cytokine, designated zalphall ligand, useful for stimulating

PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX
 XX
 PS Disclosure: Page 205-206; 256pp; English.
 XX
 CC The present sequence represents a human zalphall ligand polypeptide,
 CC which is a cytokine. The zalphall ligand is useful for stimulating the
 CC proliferation and development of haematopoietic cells in vitro and in
 CC vivo. Zalphall ligand polynucleotides can be used as primers or probes
 CC for cloning the zalphall gene. The zalphall ligand is useful for
 CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
 CC used for treating leukaemias and lymphomas. Antagonists against zalphall
 CC ligand are useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 XX
 SQ Sequence 162 AA;

Query Match 100.0%; Score 557; DB 21; Length 162;
 Best Local Similarity 100.0%; Pred. No. 3.1e-58;
 Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QLIDIVDQLKYNVDLPPEFLPAPEDVETNCWSAFSCFQKAQLKSANTGNRIINVS I 60
 DB 41 qlidivdqknyndlvpeflpapedvetncwsafscfqaqlksantgnrliinvs i 100
 QY 61 KKLKRPSTNAGRQKRLTCPCSDSYEKKPKPEFLERFKSLQ 105
 DB 101 kklkrpstnagrqrkrltcpcsdscyekppkpeflerfklsllq 145

RESULT 2

B18627
 ID B18627 standard; Protein; 519 AA.
 AC B18627;
 XX
 XX 22-JAN-2001 (first entry)
 DT
 XX
 XX Amino acid sequence of MBP-human zalphall ligand fusion in pTAP126.
 DE
 XX
 XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 KW
 XX Synthetic.
 OS Homo sapiens.
 XX
 XX WO2000053761-A2.
 PN
 XX 14-SEP-2000.
 PD
 XX
 XX 09-MAR-2000; 2000WO-US06067.
 PF
 XX
 XX 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 PI
 XX WPI; 2000-565600/52.
 DR N-PSDB; A75599.
 DR
 XX New human cytokine, designated zalphall ligand, useful for stimulating

PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX
 XX
 PS Example 31; Page 233-235; 256pp; English.
 XX
 CC The present sequence represents a MBP-human zalphall ligand fusion in
 CC the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
 CC is useful for stimulating the proliferation and development of
 CC haematopoietic cells in vitro and in vivo. Zalphall ligand
 CC polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 XX
 SQ Sequence 519 AA;

Query Match 100.0%; Score 557; DB 21; Length 519;
 Best Local Similarity 100.0%; Pred. No. 1.3e-57;
 Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QLIDIVDQLKYNVDLPPEFLPAPEDVETNCWSAFSCFQKAQLKSANTGNRIINVS I 60
 DB 398 qlidivdqknyndlvpeflpapedvetncwsafscfqaqlksantgnrliinvs i 457
 QY 61 KKLKRPSTNAGRQKRLTCPCSDSYEKKPKPEFLERFKSLQ 105
 DB 458 kklkrpstnagrqrkrltcpcsdscyekppkpeflerfklsllq 502

RESULT 3

B18624
 ID B18624 standard; Protein; 146 AA.
 AC B18624;
 XX
 XX 22-JAN-2001 (first entry)
 DT
 XX
 XX A mouse zalphall ligand polypeptide.
 DE
 XX
 XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 KW
 XX Mus musculus.
 OS
 XX WO2000053761-A2.
 PN
 XX 14-SEP-2000.
 PD
 XX
 XX 09-MAR-2000; 2000WO-US06067.
 PF
 XX
 XX 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 PI
 XX WPI; 2000-565600/52.
 DR N-PSDB; A75580.
 DR
 XX New human cytokine, designated zalphall ligand, useful for stimulating

PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX
 PS Disclosure; Page 222-223; 256pp; English.
 XX
 CC The present sequence represents a mouse zalphall ligand polypeptide,
 CC which is a cytokine. The zalphall ligand is useful for stimulating the
 CC proliferation and development of haematopoietic cells in vitro and in
 CC vivo. Zalphall ligand polynucleotides can be used as primers or probes
 CC for cloning the zalphall gene. The zalphall ligand is useful for
 CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
 CC used for treating leukaemias and lymphomas. Antagonists against zalphall
 CC ligand are useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 SQ Sequence 146 AA;

Query Match 58.3%; Score 325; DB 21; Length 146;
 Best Local Similarity 59.6%; Pred. No. 8.2e-31;
 Matches 62; Conservative 17; Mismatches 25; Indels 0; Gaps 0;

QY 2 LIDIVDLKYNVNDLVPPELPAPEDVETNCWAFSCFQAKLKSANTGNRIINVSIIK 61
 DB 35 LIDIVEGLKIYENDLDPellisapqdvkgchcehaafacfqkaklkpsnpgnktfiidlva 94
 QY 62 KLKRKPPSTNAGRQKHRLTCTPCSDSYEKPKPEFLERFKSLIQ 105
 DB 95 qlrrrlparrgqkqhhiakpcsdseyekrtpeflerikwllq 138

RESULT 4
 B18628
 ID B18628 standard; Protein; 510 AA.
 AC B18628;
 XX
 XX 22-JAN-2001 (first entry)
 XX
 DE Amino acid sequence of MBP-mouse zalphall ligand fusion in pTAP134.
 XX
 KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX
 OS Synthetic.
 OS Mus musculus.
 XX
 PN WO200053761-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06067.
 XX
 PR 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 XX
 PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 XX
 DR WPI; 2000-565600/52.
 DR N-PSDB; A75602.
 XX
 PT New human cytokine, designated zalphall ligand, useful for stimulating

PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX
 PS Example 31; Page 239-240; 256pp; English.
 XX
 CC The present sequence represents a MBP-mouse zalphall ligand fusion in
 CC the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
 CC is useful for stimulating the proliferation and development of
 CC haematopoietic cells in vitro and in vivo. Zalphall ligand
 CC polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 SQ Sequence 510 AA;

Query Match 58.3%; Score 325; DB 21; Length 510;
 Best Local Similarity 59.6%; Pred. No. 4e-30;
 Matches 62; Conservative 17; Mismatches 25; Indels 0; Gaps 0;

QY 2 LIDIVDLKYNVNDLVPPELPAPEDVETNCWAFSCFQAKLKSANTGNRIINVSIIK 61
 DB 399 LIDIVEGLKIYENDLDPellisapqdvkgchcehaafacfqkaklkpsnpgnktfiidlva 458
 QY 62 KLKRKPPSTNAGRQKHRLTCTPCSDSYEKPKPEFLERFKSLIQ 105
 DB 459 qlrrrlparrgqkqhhiakpcsdseyekrtpeflerikwllq 502

RESULT 5
 B18625
 ID B18625 standard; Peptide; 40 AA.
 AC B18625;
 XX
 XX 22-JAN-2001 (first entry)
 XX
 DE Antigeninoc peptide derived from a human zalphall ligand polypeptide.
 XX
 KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX
 OS Homo sapiens.
 XX
 PN WO200053761-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06067.
 XX
 PR 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 XX
 PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 XX
 DR WPI; 2000-565600/52.
 XX
 PT New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro

PT and in vivo, and for treating tumourigenesis -
 PS Example 34; Page 227; 256pp; English.
 XX
 CC The present sequence was used to raise antibodies, and is derived from
 CC a human zalphall ligand polypeptide. Zalphall ligand is a cytokine.
 CC The zalphall ligand is useful for stimulating the proliferation and
 CC development of haematopoietic cells in vitro and in vivo. Zalphall
 CC ligand polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 SQ Sequence 40 AA;
 Query Match 29.4%; Score 164; DB 21; Length 40;
 Best Local Similarity 100.0%; Pred. No. 1.9e-12;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY .1 QLIDIVDLKKNVNDLPFLPAPEDVETNC 31
 Db 10 qlidivdlkknvndlpflpapedvetnc 40
 RESULT 6
 BI8626
 ID BI8626 standard; Peptide; 32 AA.
 AC BI8626;
 XX
 XX 22-JAN-2001 (first entry)
 DT
 DE Antigenic peptide derived from a human zalphall ligand polypeptide.
 XX
 XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 KW
 XX Homo sapiens.
 OS
 XX WO200053761-A2.
 PN
 XX 14-SEP-2000.
 PD
 XX 09-MAR-2000; 2000WO-US06067.
 PF
 XX 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 PI
 XX WPI; 2000-565600/52.
 DR
 XX New human cytokine, designated zalphall ligand, useful for stimulating
 XX the proliferation and/or development of haematopoietic cells in vitro
 XX and in vivo, and for treating tumourigenesis -
 XX
 XX Example 34; Page 227; 256pp; English.
 PS
 XX The present sequence was used to raise antibodies, and is derived from

CC a human zalphall ligand polypeptide. zalphall ligand is a cytokine.
 CC The zalphall ligand is useful for stimulating the proliferation and
 CC development of haematopoietic cells in vitro and in vivo. Zalphall
 CC ligand polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 SQ Sequence 32 AA;
 Query Match 23.9%; Score 133; DB 21; Length 32;
 Best Local Similarity 100.0%; Pred. No. 6.8e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 82 CPSCDSYEKKPKPEFLERFKSLIQ 105
 Db 1 cpscdsyekppkeflerfkslilq 24
 RESULT 7
 Y54825
 ID Y54825 standard; Protein; 135 AA.
 XX
 AC Y54825;
 XX
 XX 04-FEB-2000 (first entry)
 DT
 DE Human Interleukin-15 protein sequence.
 XX
 DE
 KW Interleukin-15; IL-15; antisense inhibitor; untranslated region; 5'UTR;
 KW 3'UTR; human; infection; inflammation; tumour; therapy; diagnosis.
 XX
 XX Homo sapiens.
 OS
 XX US5985663-A.
 PN
 XX 16-NOV-1999.
 PD
 XX 25-NOV-1998; 98US-0200141.
 PF
 XX 25-NOV-1998; 98US-0200141.
 PR
 XX (ISIS-) ISIS PHARM INC.
 PA
 XX Bennett CF, Cowsett LM;
 PI WPI; 2000-022283/02.
 DR N-PSDB; 237358.
 DR
 XX Antisense compound useful for inhibiting human interleukin-15
 PT expression useful for treating diseases associated with interleukin-15
 PT expression -
 PT
 XX Example 13; Column 43-44; 31pp; English.
 PS
 XX This sequence is the human interleukin-15. The invention relates to
 CC antisense compounds that are targeted to a 5' or 3' untranslated region
 CC (5'UTR or 3'UTR) of a nucleic acid molecule encoding human interleukin-15
 CC (IL-15), and inhibit the expression of human IL-15. The antisense
 CC inhibitors are useful for inhibiting expression of IL-15 in human
 CC cells or tissues in vitro, for treating humans or other animals suspected
 CC of having or being prone to a disease associated with IL-15 expression,
 CC e.g. infections, inflammation or tumours. The inhibitors can also be used

The deposit was named 141-hEtF. R83435 is a mammalian mature IL-15 polypeptide. It is a generic sequence which encompasses

CC R83309 (simian) and R83310 (human) IL-15 mature polypeptides.
XX
SQ Sequence 114 AA;

```

Query Match      17.1%  Score 95.5;  DB 16;  Length 114;
Best Local Similarity 25.2%;  Pred. No. 0.00093;
Matches 29;  Conservative 22;  Mismatches 43;  Indels 21;  Gaps 5;

QY 3 IDIYDQLKNYNDLVPEE-----LPAPEDVETNCWESAFSCF-OKAQLKSANTGNN----- 52
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 nvvisdlkk-iedligsmhidatlytesdvhpscvktamkcflllelqvgisesgdxxihd 61

QY 53 --ERIIINVISIKKLARKPPSTNAGRQRKHRLTCPSCDSYKXKPPPEFIERFVSLIQ 105
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 62 tvenlilannxlssngnxltseq-----ckeceeleeknikelfqsfvhiwq 108

```

RESULT	10
R66928	
ID	R66928 standard; Protein; I14 AA.
XX	
AC	R66928;
XX	
DT	04-SEP-1995 (first entry)
XX	
DE	Mammalian IL-15.
XX	
KW	Interleukin-15; IL-15; sIL-15; T-
KW	antitumor; virucide.
XX	
OS	Mammalia.

PH	Key	Location/Qualifiers
FT	Misc-difference 52	/label= Leu, His
FT	Misc-difference 57	/label= Ala, Thr
FT	Misc-difference 58	/label= Ser, Asp
FT	Misc-difference 73	/label= Ser, Ile
FT	Misc-difference 80	/label= Val, Ile

XX	2A9402636-A.	
PN		
XX		
XX	28-DEC-1994.	
PD		
XX		
XX	18-APR-1994; 94ZA-0002636.	
PF		
XX		
XX	18-APR-1994; 94ZA-0002636.	
PR		
XX		
XX	(IMMV) IMMUNEX CORP.	
PA		
XX		
XX	Anderson DM, Eisenman JR, Fung V, Grabstein KH;	
PI	Rauch C;	
PI		

WI; 1995-082473/11.

New purified interleukin-15 - which induces T cell proliferation and differentiation, used for the treatment of tumours and viral infection

Claim 1; Page 33; 47pp; English.

Simian and human IL-15 cDNAs (084583-84) can be used to obtain cDNAs encoding other mammalian homologs of IL-15. A general sequence of mammalian IL-15 is claimed.

AA	Sequence	114 AA;
----	----------	---------

Query Match	17.1%	Score 95.5;	DB 16;	Length 114;
Best Local Similarity	25.2%	Pred. NO. 0.00093;		
Matches 29; Conservative	22;	Mismatches 43;	Indels 21;	Gaps 5;

3 IDIVDOLKNVNDLVEF-----LPAPEDVETNCWSAFSCF-QKAQLKSAKTGNN----- 52
 y y
 b b
 3 3
 y y
 b b
 53 62 62
 y y
 b b

RESULT 11
09099
D W09099 standard; Protein; 114 AA.

A	W09099;
C	11-MAR-1997 (first entry)
X	Simian mature epithelium derived T cell factor.
X	sF7F; African green monkey; CV1/EBNA cell; T-cell; B-cell;
T	lymphocyte; proliferation; differentiation; gastrointestinal;
X	HIV infection; human immunodeficiency virus.
X	Cercopithecus aethiops.
S	

Accession	Key	Location/Qualifiers
U00096	Protein	1..114
U00096		/label= mature SETD

US5574138-A.
12-NOV-1996.

X
F
08-MAR-1993; 93US-0031399.

A R 22-FEB-1995; 95US-0393305.

R 22-APR-1994; 94US-0233606.

A (IMMV) IMMUNEX CORP.

I Anderson DM, Eisenman JR, Fung V, Grabstein KH;
I Rauch C;

WPI; 1996-517923/51.
N-PSDB: T49455.

New epithelium derived T cell factor - induces proliferation of T and B cells, stimulates destruction of tumour and virus-infected cells and protects against toxicity, partic. for treating intestinal disease and HIV infection

Claim 1; Fig. 1; 35pp; English.

The simian ETF (epithelium derived T cell factor) was isolated from African green monkey CV1/EBNA cell conditioned medium. The N-terminal sequence of the purified sETF was determined and then PCR primers were designed based on the sequence information. A 92 bp fragment was amplified from CV1/EBNA DNA and was used as a probe to screen a CV1/EBNA cDNA library for the full-length sETF coding sequence. Mature sETF induces proliferation and/or differentiation of precursor or mature T cells and is useful for promoting long term in vitro culture of T-lymphocytes and T-cell lines. It is used for treating gastrointestinal diseases including peptic ulcer, colitis and malignancy and for treating HIV infection.

A

Q Sequence 114 AA;

Best Local Similarity	25.2%;	Pred. No. 0.00093;	
Matches	29; Conservative	22; Mismatches	43; Indels
			21; Gaps
QY	3	IDIVDOLKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-QKAOCLKSANTGNN----	52
Db	3	vnvisdLkK-iedlqsmhdIatlytesdVhpckvtamKcflllelqvixesgdxxlhd	61
QY	53	--ERIINVSIIKKLRPPSTNAGRRQKHRLTCPSCDSEYKPKPEFLERFKSLQ	105
Db	62	tvenlilannxissngntesg-----ckeceeleeknikelfqsfvhiq	108
RESULT	13		
W39186	ID	W39186 standard; Protein; 114 AA.	
XX	AC	W39186;	
AC	XX		
XX	XX	08-MAY-1998 (first entry)	
XX	XX	Simian epithelium derived T-cell factor mature protein.	
DE	XX		
XX	XX	Epithelium derived T-cell factor; ETF; simian; gastrointestinal disease;	
KW	XX	B-cell stimulation; epithelium; villi; CD4+; CD8+; T-cell stimulation;	
KW	XX	treatment; prevention.	
XX	XX		
XX	OS	Simian.	
XX	Key	Location/Qualifiers	
FX	Protein	1..114	
FT	FT	/label= ETF	
XX	XX	US5707616-A.	
PN	XX		
PD	XX	13-JAN-1998.	
PF	XX	04-OCT-1996; 96US-0726817.	
XX	XX		
XX	XX	22-FEB-1995; 95US-0393305.	
PR	XX	08-MAR-1993; 93US-0031399.	
PR	XX	22-APR-1994; 94US-0233606.	
XX	XX	(IMMV) IMMUNEX CORP.	
XX	XX		
PI	XX	Anderson DM, Eisenman JR, Fung V, Grabstein KH;	
PI	XX	Rauch C;	
XX	XX		
DR	XX	WPI; 1998-100295/09.	
DR	XX	N-PSDB; V02873.	
XX	XX		
PT	XX	Treatment or prevention of gastrointestinal diseases - by	
PT	XX	administering epithelium-derived T-cell factor polypeptide	
XX	XX		
PS	XX	Claim 1a; Column 37-38; 34pp; English.	
XX	XX		
XX	XX	This sequence represents a simian epithelium-derived T-cell factor (ETF)	
CC	XX	mature protein which is used in a method for treating or preventing	
CC	XX	gastrointestinal disease. These polypeptides have particular application	
CC	XX	in the treatment of gastrointestinal disorders associated with disruption	
CC	XX	of the gastrointestinal epithelium or villi such as chemotherapy- and	
CC	XX	radiation-therapy induced enteritis (gut toxicity), mucositis, peptic	
CC	XX	ulcer disease, gastroenteritis and colitis, villus atrophic disorders,	
CC	XX	malignancy and inflammatory bowel disease. ETF polypeptides may also be	
CC	XX	useful in the treatment of human immunodeficiency virus (HIV) and	
CC	XX	HIV-associated disease due to their ability to stimulate CD4+ and CD8+	
CC	XX	cells. Biologically active ETF may be used to treat a variety of other	
CC	XX	diseases or conditions where T-cell or B cell stimulation is desired.	
XX	XX	Sequence	114 AA;
			SQ

Query Match 17.1%; Score 95.5; DB 19; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00093;

Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;

QY 3 IDIVDQKNYVNDLVPF-----LPAPEDVETNCWESAFSCF-----QKAQLKSANTGNNE 53
 Db 3 vnvissdlkk-iedllqsmhidatlytesdvhpckvtamckfllleqlvislesgdsahd 61

QY 54 RIINVSIRKLRKPPSTNAGRQKRLTCPCSDSYEKPPKPFLEFRKSLQ 105
 Db 62 tvenlii--lannslssngvntes---gckeceeleeknikelflsfvhivq 108

RESULT 14

Y03757

ID Y03757 standard; Protein; 114 AA.

XX AC Y03757;

XX DT 10-JUN-1999 (first entry)

XX DE Simian epithelium-derived T-cell factor (ETF) mature protein sequence.

XX KW Epithelium-derived T-cell factor; ETF; immunoassay; enteritis; ulcer;
 KW T cell proliferation; gastrointestinal disease; mucositis; colitis;
 KW gastroenteritis; villus atrophy; malignancy; inflammatory bowel disease;
 KW human immune deficiency virus; tumour; simian.

XX OS Mammalia.

XX PN US5892001-A.

XX PD 06-APR-1999.

XX PF 04-OCT-1996; 96US-0725969.

XX PR 22-FEB-1995; 95US-0393305.

XX PR 08-MAR-1993; 93US-0031399.

XX PR 22-APR-1994; 94US-0233606.

XX PR 04-OCT-1996; 96US-0725969.

XX PA (IMMV) IMMUNEX CORP.

XX PI Anderson DM, Eisenman JR, Fung V, Grabstein KH;

XX PI Rauch C;

XX DR WPI; 1999-253930/21.

XX DR N-PSDB; X29479.

XX PT Antibodies specific for epithelium-derived T-cell growth factor

XX PS Claim 1; Fig 1; 34pp; English.

XX CC The invention relates to an isolated antibody that binds specifically to
 CC a simian or human epithelium-derived T-cell factor (ETF) polypeptide. The
 CC antibodies are used, optionally when immobilized or labeled, to detect
 CC and quantify ETF in standard immunoassays. They may also be used as
 CC diagnostic and therapeutic agents, e.g. when conjugated to toxins (or
 CC their precursors) or radionuclides. ETF induces proliferation and/or
 CC establishing long term in vitro cultures; and is also used to treat
 CC gastrointestinal disease (e.g. enteritis or mucositis induced by
 CC chemotherapy or radiation, peptide ulcer, gastroenteritis, colitis,
 CC villus atrophy, malignancy and inflammatory bowel disease), to treat
 CC human immune deficiency virus infection or associated disease, or
 CC generally in any situation requiring stimulation of T or B cell
 CC proliferation, secretion of immunoglobulins or certain cytokines.
 CC increased anti-infectious disease immunity, induction of T-cell lytic
 CC activity or increased destruction of tumour or virus-infected cells. The
 CC present sequence represents the mature active sequence of simian ETF
 CC polypeptide.

XX Sequence 114 AA;

Query Match 17.1%; Score 95.5; DB 20; Length 114;
 Best Local Similarity 25.9%; Pred. No. 0.00093;
 Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;

QY 3 IDIVDQKNYVNDLVPF-----LPAPEDVETNCWESAFSCF-----QKAQLKSANTGNNE 53
 Db 3 vnvissdlkk-iedllqsmhidatlytesdvhpckvtamckfllleqlvislesgdsahd 61

QY 54 RIINVSIRKLRKPPSTNAGRQKRLTCPCSDSYEKPPKPFLEFRKSLQ 105
 Db 62 tvenlii--lannslssngvntes---gckeceeleeknikelflsfvhivq 108

RESULT 15

Y52309

ID Y52309 standard; Protein; 114 AA.

XX AC Y52309;

XX DT 09-FEB-2000 (first entry)

XX DE Mature simian epithelium-derived T-cell factor (ETF).

XX KW ETF; epithelium-derived T-cell factor; T-cell; T-lymphocyte;
 KW proliferation; differentiation; growth factor; precursor; mature; CD4+;
 KW CD8+; gastrointestinal; epithelium; cell culture; treatment; HIV;
 KW gastrointestinal disease; gastroenteritis; colitis;
 KW inflammatory bowel disease; villus atrophic disorder; enteritis;
 KW chemotherapy; radiotherapy; gut toxicity; cancer; side effect;
 KW tolerated dose.

XX OS Mammalia.

XX PN US5985262-A.

XX PD 16-NOV-1999.

XX PF 03-FEB-1997; 97US-0794524.

XX PR 22-FEB-1995; 95US-0393305.

XX PR 04-OCT-1996; 96US-0726817.

XX PR 08-MAR-1993; 93US-0031399.

XX PR 22-APR-1994; 94US-0233606.

XX PA (IMMV) IMMUNEX CORP.

XX PI Rauch C, Fung V, Eisenman JR, Grabstein KH, Anderson DM;

XX PI WPI; 2000-022267/02.

XX DR N-PSDB; Z38244.

XX PT Stimulation of T-cells in human immunodeficiency virus infected
 PT patients -

XX PS Claim 1; Fig 1; 33pp; English.

XX CC This sequence represents mature simian epithelium-derived T-cell factor
 CC (ETF). This is a previously unidentified T-cell growth factor which
 CC stimulates precursor or mature CD4+ and CD8+ T-lymphocytes to proliferate
 CC and differentiate. It also promotes proliferation of the gastrointestinal
 CC epithelium. The protein can be used to promote long-term in vitro culture
 CC of T-lymphocytes and T-cell lines. ETF can be used for treating HIV
 CC infection, HIV-associated diseases, and other diseases or conditions
 CC where stimulation of T-cell proliferation would be desirable e.g., it
 CC could be used to augment the destruction of tumour cells or virally-
 CC infected cells. ETF may also be used to treat or prevent gastrointestinal
 CC disease, including chemotherapy and radiotherapy associated enteritis,
 CC gastroenteritis, colitis, inflammatory bowel disease and villus atrophic
 CC disorders. Chemotherapy and radiotherapy associated enteritis (gut
 CC toxicity) results in bleeding and sepsis due to gastrointestinal flora
 CC entering the blood, and thus can limit the dosage of therapeutic agent
 CC administered to a cancer patient. ETF may therefore be used to increase
 CC the tolerated doses radiotherapy and chemotherapy.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:09:41 ; Search time 58.85 Seconds
(without alignments)
34.276 Million cell updates/sec

Title: US-09-522-217-2_COPY_41_145

Perfect score: 557
Sequence: 1 QLIDVQLKYNVDLPEF.....DSYEKKPKFLEPKSLIQ 105

Scoring table: BLASUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95.5	17.1	114	1 US-08-031-399-6	Sequence 6, Appli
2	95.5	17.1	114	1 US-08-031-399-12	Sequence 12, Appli
3	95.5	17.1	114	1 US-08-393-305-3	Sequence 3, Appli
4	95.5	17.1	114	1 US-08-726-817-3	Sequence 3, Appli
5	95.5	17.1	114	1 US-08-504-042-6	Sequence 6, Appli
6	95.5	17.1	114	1 US-08-504-042-12	Sequence 12, Appli
7	95.5	17.1	114	2 US-08-725-969-3	Sequence 3, Appli
8	95.5	17.1	114	2 US-08-794-524-3	Sequence 3, Appli
9	95.5	17.1	114	4 US-09-189-193-3	Sequence 3, Appli
10	95.5	17.1	114	5 PCT-US94-03793-6	Sequence 6, Appli
11	95.5	17.1	114	5 PCT-US94-03793-12	Sequence 12, Appli
12	95.5	17.1	162	1 US-08-031-399-5	Sequence 5, Appli
13	95.5	17.1	162	1 US-08-393-305-2	Sequence 2, Appli
14	95.5	17.1	162	1 US-08-535-733-2	Sequence 2, Appli
15	95.5	17.1	162	1 US-08-726-817-2	Sequence 2, Appli
16	95.5	17.1	162	1 US-08-504-042-5	Sequence 5, Appli
17	95.5	17.1	162	2 US-08-725-969-2	Sequence 2, Appli
18	95.5	17.1	162	2 US-08-794-524-2	Sequence 2, Appli
19	95.5	17.1	162	3 US-08-842-947-6	Sequence 6, Appli
20	95.5	17.1	162	4 US-09-189-193-2	Sequence 2, Appli
21	95.5	17.1	162	5 PCT-US94-03793-5	Sequence 5, Appli
22	95.5	17.1	162	5 PCT-US96-06423-2	Sequence 2, Appli
23	90.5	16.2	114	1 US-08-031-399-3	Sequence 3, Appli
24	90.5	16.2	114	1 US-08-393-305-6	Sequence 6, Appli
25	90.5	16.2	114	1 US-08-726-817-6	Sequence 3, Appli
26	90.5	16.2	114	1 US-08-504-042-3	Sequence 3, Appli
27	90.5	16.2	114	2 US-08-725-969-6	Sequence 6, Appli

28	90.5	16.2	114	2 US-08-794-524-6	Sequence 6, Appli
29	90.5	16.2	114	4 US-09-189-193-6	Sequence 6, Appli
30	90.5	16.2	114	5 PCT-US94-03793-3	Sequence 3, Appli
31	90.5	16.2	122	1 US-08-300-903A-3	Sequence 3, Appli
32	90.5	16.2	162	1 US-08-031-399-2	Sequence 2, Appli
33	90.5	16.2	162	1 US-08-393-305-5	Sequence 5, Appli
34	90.5	16.2	162	1 US-08-284-393B-9	Sequence 9, Appli
35	90.5	16.2	162	1 US-08-726-817-5	Sequence 5, Appli
36	90.5	16.2	162	2 US-08-504-042-2	Sequence 2, Appli
37	90.5	16.2	162	2 US-08-725-969-5	Sequence 5, Appli
38	90.5	16.2	162	2 US-08-794-524-5	Sequence 5, Appli
39	90.5	16.2	162	3 US-08-842-947-8	Sequence 8, Appli
40	90.5	16.2	162	4 US-09-189-193-5	Sequence 5, Appli
41	90.5	16.2	162	5 PCT-US94-03793-2	Sequence 2, Appli
42	90.5	16.2	162	5 PCT-US95-08950-9	Sequence 9, Appli
43	82	14.7	337	1 US-08-442-043A-18	Sequence 18, Appli
44	69	12.4	1220	3 US-08-930-966A-2	Sequence 2, Appli
45	66.5	11.9	3665	2 US-08-222-617A-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1
US-08-031-399-6
; Sequence 6, Application US/08031399
; Patent No. 5552303
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Epithelium-derived T-cell Factor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/031,399
; FILING DATE: 19930308
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-031-399-6

Query Match 17.1%; Score 95.5; DB 1; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00024;
Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;
Oy 3 IDIVIDQLKYNVDLPEF-----LPAPEDVTNCEWSAFSCF----QKAQLKSANTGNNE 53
::: || : ||: | || : | || : |

Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTAMKCFLELQVLSLESGDASIH 61
QY 54 RIINVSIIKKLRKPPSTNAGRQKRLHLCPCSDSYEKPKPKPELRFKSLQ 105
Db 62 TVENLII--LANNSSNGNVTES--GCKECEELEEKNIKEFLQSFVHIQ 108

RESULT 2
US-08-031-399-12
; Sequence 12, Application US/08031399
; Patent No. 5522303
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Epithelium-derived T-cell Factor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/031,399
; FILING DATE: 19930308
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-031-399-12

Query Match 17.1%; Score 95.5; DB 1; Length 114;
Best Local Similarity 25.2%; Pred. No. 0.00024;
Matches 29; Conservative 22; Mismatches 43; Indels 21; Gaps 5;
QY 3 IDIVOLKNYNDLVPEF-----LPAPEDVTNCWSAFSCF--OKAQLKSANTGNN----- 52
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTAMKCFLELQVLSLESGDASIH 61
QY 53 --ERLIINSIKKLRKPPSTNAGRQKRLHLCPCSDSYEKPKPKPELRFKSLQ 105
Db 62 TVENLIIANNLSSNGNXTESG-----CKECEELEEKNIKEFLQSFVHIQ 108

RESULT 3
US-08-393-305-3
; Sequence 3, Application US/08393305
; Patent No. 5574138
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor

; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,305
; FILING DATE: 22-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-393-305-3

Query Match 17.1%; Score 95.5; DB 1; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00024;
Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;
QY 3 IDIVOLKNYNDLVPEF-----LPAPEDVTNCWSAFSCF--OKAQLKSANTGNE 53
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTAMKCFLELQVLSLESGDASIH 61
QY 54 RIINVSIIKKLRKPPSTNAGRQKRLHLCPCSDSYEKPKPKPELRFKSLQ 105
Db 62 TVENLII--LANNSSNGNVTES--GCKECEELEEKNIKEFLQSFVHIQ 108

RESULT 4
US-08-726-817-3
; Sequence 3, Application US/08726817
; Patent No. 5707616
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,817

; LENGTH: 114 amino acids

$\frac{1}{2}$ $\frac{1}{3}$ $\frac{1}{4}$ $\frac{1}{5}$ $\frac{1}{6}$ $\frac{1}{7}$ $\frac{1}{8}$ $\frac{1}{9}$ $\frac{1}{10}$ $\frac{1}{11}$ $\frac{1}{12}$ $\frac{1}{13}$ $\frac{1}{14}$ $\frac{1}{15}$ $\frac{1}{16}$ $\frac{1}{17}$ $\frac{1}{18}$ $\frac{1}{19}$ $\frac{1}{20}$ $\frac{1}{21}$ $\frac{1}{22}$ $\frac{1}{23}$ $\frac{1}{24}$ $\frac{1}{25}$ $\frac{1}{26}$ $\frac{1}{27}$ $\frac{1}{28}$ $\frac{1}{29}$ $\frac{1}{30}$ $\frac{1}{31}$ $\frac{1}{32}$ $\frac{1}{33}$ $\frac{1}{34}$ $\frac{1}{35}$ $\frac{1}{36}$ $\frac{1}{37}$ $\frac{1}{38}$ $\frac{1}{39}$ $\frac{1}{40}$ $\frac{1}{41}$ $\frac{1}{42}$ $\frac{1}{43}$ $\frac{1}{44}$ $\frac{1}{45}$ $\frac{1}{46}$ $\frac{1}{47}$ $\frac{1}{48}$ $\frac{1}{49}$ $\frac{1}{50}$ $\frac{1}{51}$ $\frac{1}{52}$ $\frac{1}{53}$ $\frac{1}{54}$ $\frac{1}{55}$ $\frac{1}{56}$ $\frac{1}{57}$ $\frac{1}{58}$ $\frac{1}{59}$ $\frac{1}{60}$ $\frac{1}{61}$ $\frac{1}{62}$ $\frac{1}{63}$ $\frac{1}{64}$ $\frac{1}{65}$ $\frac{1}{66}$ $\frac{1}{67}$ $\frac{1}{68}$ $\frac{1}{69}$ $\frac{1}{70}$ $\frac{1}{71}$ $\frac{1}{72}$ $\frac{1}{73}$ $\frac{1}{74}$ $\frac{1}{75}$ $\frac{1}{76}$ $\frac{1}{77}$ $\frac{1}{78}$ $\frac{1}{79}$ $\frac{1}{80}$ $\frac{1}{81}$ $\frac{1}{82}$ $\frac{1}{83}$ $\frac{1}{84}$ $\frac{1}{85}$ $\frac{1}{86}$ $\frac{1}{87}$ $\frac{1}{88}$ $\frac{1}{89}$ $\frac{1}{90}$ $\frac{1}{91}$ $\frac{1}{92}$ $\frac{1}{93}$ $\frac{1}{94}$ $\frac{1}{95}$ $\frac{1}{96}$ $\frac{1}{97}$ $\frac{1}{98}$ $\frac{1}{99}$ $\frac{1}{100}$

```

Db      3  VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVTKAMKCFLELQVIXESGDXIHD 61
QY      53  --ERINVISIKLKKPSTWAGRRQKRLTCTPSCSYEKKKPKFELEFKSLQ 105
Db      62  TVENLIILANNXLSNGNXTSG-----CXEELEEKNIKEQLQSFVHQ 108

```

```

RESULT      7
US-08-725-969-3
; Sequence 3, Application US/08725969
; Patent No. 5892001
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104

```

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/725,969
FILING DATE: 04-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/393,305
FILING DATE: 22-FEB-1995
ATTORNEY/AGENT INFORMATION:

```

NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 480052.409C2
TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;

```

Query Match	17.1%	Score 95.5;	DB 2;	Length 114;
Best Local Similarity	25.9%	Pred. No. 0.00024;		
Matches 29;	Conservative	22;	Mismatches 46;	Indels 15;
			Gaps	5;

Qy 3 IDIVDLKYNVNDLVPEF-----LPAPEDVETNCWSAFSCF---KQAQLKSANTGNNE 53

Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTKMCFLELQVVISLESGDASIH 61

Qy	54 RIINVSIKKLKRPPSTNACRRQKHRLTCPSCDSYEKKPKPEFLERFKSLIQ 105
Db	62 TVENLII--LANSSLSSNGNVTES--GCKECEEELEERNIKEFLOSFVHIQV 108

RESULT 8
US-08-794-524-3
; Sequence 3, Application US/08794524
; Patent No. 5985262
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk

APPLICANT: Eisenman, June
 APPLICANT: Fung, Victor
 APPLICANT: Rauch, Charles
 TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Seed and Berry
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/794,524
 FILING DATE: 03-FEB-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/393,305
 FILING DATE: 22-FEB-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: McMasters, David D.
 REGISTRATION NUMBER: 33,963
 REFERENCE/DOCKET NUMBER: 480052.409C2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 206-622-4900
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 114 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-794-524-3

Query Match	17.1%	Score 95.5;	DB 2;	Length 114;
Best Local Similarity	25.9%;	Pred. No. 0.00024;		
Matches 29;	Conservative 22;	Mismatches 46;	Indels 15;	Gaps 5;

```

QY 3 IDIVOLKNYNDLPEF-----LPAPDEVTCNWSAFSCF-----QKAOUKSANTGNE 53
      ::::: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 3 VNVISDLAK--IEDLIQSMDIATLYTESDVPSCVKVTAMKFLJELQVLISGDSATHD 61

QY 54 RIINVSIKKLKRRKPPSTNAGRQRKRLHLCPCSDSYEKKPPKEFERFKSLQ 105
      ::::: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 62 TVENLII--LANNSSNGVNTES---CKCEEELEEKNIKEFLOSQFVHIQ 108
      ::::: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

```

RESULT 9
US-09-189-193-3
Sequence 3, Application US/09189193
Patent No. 6184359
GENERAL INFORMATION:
APPLICANT: Grabstein, Kenneth
APPLICANT: Anderson, Dirk
APPLICANT: Eisenman, June
APPLICANT: Fung, Victor
APPLICANT: Rauch, Charles
TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk


```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/189,193
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,305
; FILING DATE: 22-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McMAsters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-189-193-3

```

```

Query Match 17.1%; Score 95.5; DB 4; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00024;
Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;

QY 3 IDVDQLKNVNDLVPEF-----LPAPEDVTNCENSAFSCF-----OKAQLKSANTGNNE 53
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVLSGSDASHD 61
QY 54 RIINVSIKKLKRPPSTNAGRQKHRLTQPCSDSYEKKPKPELRFKSLIQ 105
Db 62 TVENLII--LANNLSNGNVTES---GCKECELEEKNIKEFLQSFVHIQ 108

```

```

RESULT 10
PCT-US94-03793-6
; Sequence 6, Application PC/TUS9403793
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Interleukin-15
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03793
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; PCT-US94-03793-6

```

```

Query Match 17.1%; Score 95.5; DB 5; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00024;
Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;

QY 3 IDVDQLKNVNDLVPEF-----LPAPEDVTNCENSAFSCF-----OKAQLKSANTGNNE 53
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVLSGSDASHD 61
QY 54 RIINVSIKKLKRPPSTNAGRQKHRLTQPCSDSYEKKPKPELRFKSLIQ 105
Db 62 TVENLII--LANNLSNGNVTES---GCKECELEEKNIKEFLQSFVHIQ 108

```

```

RESULT 11
PCT-US94-03793-12
; Sequence 12, Application PC/TUS9403793
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Interleukin-15
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03793
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; PCT-US94-03793-12

```

```

Query Match 17.1%; Score 95.5; DB 5; Length 114;
Best Local Similarity 25.2%; Pred. No. 0.00024;
Matches 29; Conservative 22; Mismatches 43; Indels 21; Gaps 5;

QY 3 IDVDQLKNVNDLVPEF-----LPAPEDVTNCENSAFSCF-----OKAQLKSANTGNNE 52
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVLSGSDASHD 61
QY 53 --ERIINVSIKKLKRPPSTNAGRQKHRLTQPCSDSYEKKPKPELRFKSLIQ 105

```

; CORRESPONDENCE ADDRESS:

REFERENCE/DOCKET NUMBER: 2833

TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-587-0430
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-535-733-2

Query Match 17.1%; Score 95.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.00038;
Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;

QY 3 IDIVDLKKNYNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKAQLKSANTGNNE 53
Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVLSLESGDASIHD 109
QY 54 RIINVISIKLKRKPPSTNAGRQKHRLTCPCSDSYEKKPPKEFLERFKSLQ 105
Db 110 TVENLII--LANNSSLSNGNVTES---GCKECELEEKNIKEFLQSFVHIVQ 156

RESULT 15

US-08-726-817-2
Sequence 2, Application US/08726817
Patent No. 5707616
GENERAL INFORMATION:
APPLICANT: Grabstein, Kenneth
APPLICANT: Anderson, Dirk
APPLICANT: Eisenman, June
APPLICANT: Fung, Victor
APPLICANT: Rauch, Charles
TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,817
FILING DATE: 04-OCT-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/393,305
FILING DATE: 22-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: McMasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 480052.409C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-726-817-2

Query Match 17.1%; Score 95.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.00038;
Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;

QY 3 IDIVDLKKNYNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKAQLKSANTGNNE 53
Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVLSLESGDASIHD 109
QY 54 RIINVISIKLKRKPPSTNAGRQKHRLTCPCSDSYEKKPPKEFLERFKSLQ 105
Db 110 TVENLII--LANNSSLSNGNVTES---GCKECELEEKNIKEFLQSFVHIVQ 156

Search completed: May 23, 2001, 11:09:42
Job time: 65 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:12:49 ; Search time 70.54 Seconds
(without alignments)
102.295 Million cell updates/sec

Title: US-09-522-217-2_COPY_41_145
Perfect score: 557
Sequence: 1 QLIDVQLKKNVNDLVPEF.....DSYEKKPKPEFLERFKSLIQ 105

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_67.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	90.5	16.2	162	1 A53484	interleukin-15 pre
2	82	14.7	567	2 S29498	lymphocyte antigen
3	77.5	13.9	607	1 ABXL72	74K albumin precu
4	74	13.3	2197	2 B71600	variant-specific s
5	73.5	13.2	868	2 T31527	hypothetical prote
6	72	12.9	304	2 A32108	translation initia
7	71.5	12.8	195	2 S42022	ureidoglycolate hy
8	71	12.7	1101	2 T26919	hypothetical prote
9	70.5	12.7	848	2 T00372	hypothetical prote
10	70	12.6	262	2 F72858	probable methyl tr
11	70	12.6	262	2 T41813	AcMNPV orf69 - Bom
12	69.5	12.5	162	2 I49124	interleukin-15 - m
13	69	12.4	206	2 S49882	hypothetical prote
14	69	12.4	1220	2 T06403	resistance complex
15	68.5	12.3	336	2 S42632	Fit-15 protein pre
16	68.5	12.3	1257	2 T01020	hypothetical prote
17	68	12.2	329	2 C69483	hypothetical prote
18	68	12.2	397	2 A33880	syndecan 2 - huma
19	67.5	12.1	405	2 T28957	hypothetical prote
20	66.5	11.9	420	2 A25876	vitellogenin III p
21	66.5	11.9	636	2 T37941	conserved hypotet
22	66.5	11.9	1750	2 H64403	ribonucleoside-tri
23	66.5	11.9	3712	1 YGC5VC	alpha-aminoadipyl-
24	66	11.8	805	2 A46266	aryl hydrocarbon r
25	66	11.8	2924	2 T18378	variant-specific s
26	65.5	11.8	189	2 S11632	myosin regulatory
27	65.5	11.8	219	2 B84079	hypothetical prote
28	65.5	11.8	405	2 A72383	sensor histidine k
29	65.5	11.8	576	2 S50113	coillin, p80 - huma

30	65	11.7	257	2 C69230	conserved hypothet
31	65	11.7	447	2 T40423	hypothetical prote
32	65	11.7	695	2 S37439	transketolase (EC
33	65	11.7	1388	2 S70633	serine/threonine-s
34	65	11.7	1388	2 S74245	serine/threonine-s
35	64.5	11.6	259	2 S32898	hypothetical prote
36	64.5	11.6	465	2 A56679	probable RNA-bindl
37	64.5	11.6	741	2 B69400	molybdopter in oxid
38	64.5	11.6	1671	2 S71628	sensory transducti
39	64.5	11.6	4385	2 T29042	hypothetical prote
40	64	11.5	249	2 B69343	conserved hypothet
41	64	11.5	299	2 D81384	probable lipoprote
42	64	11.5	333	2 T28415	ORF MSV254 leucine
43	64	11.5	606	2 T21263	hypothetical prote
44	64	11.5	628	2 S61160	hypothetical prote
45	64	11.5	921	2 H69643	isoleucine--tRNA 1

ALIGNMENTS

RESULT 1

A53484

interleukin-15 precursor - green monkey

C:Species: Cercopithecus aethiops (green monkey, grivet)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A53484

R:Grabstein, K.H.; Eisenman, J.; Shanebeck, K.; Rauch, C.; Srinivasan, S.; Fung, V.;

Giri, J.G.

Science 264, 965-968, 1994

A:Title: Cloning of a T cell growth factor that interacts with the beta chain of the

A:Reference number: A53484; MUID:94233380

A:Accession: A53484

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-162 <GRA>

A:Cross-references: GB:003099; NID:g493521; PIDN:AAAL8416.1; PID:g493522

A:Note: the complete translation is not shown

C:Superfamily: interleukin-15

C:Keywords: growth factor

F:43-162/Product: interleukin-15 #status predicted <MAT>

F:133-133,90-136/Disulfide bonds: #status predicted

Query Match 16.2%; Score 90.5; DB 1; Length 162; .

Best Local Similarity 25.9%; Pred. No. 0.049; 46; Indels 15; Gaps 5;

Matches 29; Conservative 22; Mismatches 22

Qy 3 IDIVDLKKNVNDLVPEF-----LPAPEDVTNCWSAFSCF----QKAQLKSANTGNNE 53

Db 51 VNVISDLAK-IEDLIQSMHIDATLYTESDVHPSCKVTAMKCFLELEQLVISHESGDTDIHD 109

Qy 54 RIINVSITKKLRKPPSTNAGRQKRLRCPCSDSYEKPKPEFLERFKSLIQ 105

Db 110 TVENLI--LANNILSSNGNITES---GCKECELEEKNIKEFLQSFVHIQV 156

RESULT 2

S29498

lymphocyte antigen Ly84 precursor - mouse

N:Alternate names: 38.5K T1 glycoprotein; ST2L protein

C:Species: Mus musculus (house mouse)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Dec-2000

C:Accession: S29498; A33541; S17657; S07054

R:Yanagisawa, K.; Takagi, T.; Tsukamoto, T.; Tetsuka, T.; Tominaga, S.

FEBS Lett. 318, 83-87, 1993

A:Title: Presence of a novel primary response gene ST2L, encoding a product highly si

A:Reference number: S29498; MUID:93170492

A:Accession: S29498

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-567 <YAN>

A:Cross-references: EMBL:D13695; NID:g286100; PIDN:BAAO2854.1; PID:g286101

R:Klemen, R.; Hoffmann, S.; Werenskiold, A.K.
 Proc. Natl. Acad. Sci. U.S.A. 86, 5708-5712, 1989
 A:Title: Serum- and oncoprotein-mediated induction of a gene with sequence similarity to
 A:Reference number: A33541; MUID:89345536
 A:Accession: A33541
 A:Molecule type: mRNA
 A:Residues: 1-191, 'V', 193-328, 'SKECPSHIA' <KLE>
 R:Tominaga, S.I.; Jenkins, N.A.; Gilbert, D.J.; Copeland, N.G.; Tetsuka, T.
 Biochim. Biophys. Acta 1090, 1-8, 1991
 A:Title: Molecular cloning of the murine ST2 gene. Characterization and chromosomal map
 A:Reference number: S17657; MUID:91355215
 A:Accession: S17657
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-191, 'V', 193-328, 'SKECPSHIA' <TOM>
 A:Cross-references: EMBL:X60184; NID:g54200; PIDN:CAA42742.1; PID:g54201
 R:Tominaga, S.I.
 FEBS Lett. 258, 301-304, 1989
 A:Title: A putative protein of a growth specific cDNA from BALB/c-3T3 cells is highly si
 A:Reference number: S07054; MUID:90092495
 A:Accession: S07054
 A:Molecule type: mRNA
 A:Residues: 1-328, 'SKECPSHIA' <TO2>
 A:Cross-references: EMBL:Y07519; NID:g55517; PIDN:CAA68812.1; PID:g55518
 A:Note: it is uncertain whether Met-1, Met-7 or Met-19 is the initiator
 C:Genetics:
 A:Gene: ST2
 A:Map position: 1
 A:Introns: 27/1; 97/2; 155/3; 210/1; 233/1; 280/2
 C:Superfamily: interleukin-1 receptor type I
 C:Keywords: glycoprotein
 F:1-26/Domain: signal sequence #status predicted <SIG>
 F:27-337/Product: ST2 protein #status predicted <MAT>
 F:60.101.107.146.176.194.225.259.278/Binding site: carbohydrate (Asn) (covalent) #status

Query Match 14.7%; Score 82; DB 2; Length 567;
 Best Local Similarity 27.1%; Pred. No. 1.3;
 Matches 29; Conservative 17; Mismatches 29; Indels 32; Gaps 6;
 QY 7 DOLKNVNDLVPEFLPAPDEVETNCSEAFSCFQKALKSANTGNNGRIINVSIIKLRK 66
 Db 76 DRK-----FLPARVE-----DSGIYV-----IRSPNL-NKTGYNVTHIK---K 113
 QY 67 PPSTN-----AGRRQKRLTQPCSDSYEKPKPEFLERFKSL 103
 Db 114 PPSCTNPDLMYSTVRGSKFKITCPTIDLYNTAPVQWFKCKAL 160
 RESULT 3
 ABXL72
 74K albumin precursor - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
 C:Accession: B41682; S02693; A05288
 R:Moskatis, J.E.; Sargent, T.D.; Smith Jr., L.H.; Pastori, R.L.; Schoenberg, D.R.
 Mol. Endocrinol. 3, 464-473, 1989
 A:Title: Xenopus laevis serum albumin: sequence of the complementary deoxyribonucleic ac
 e during development.
 A:Reference number: A41682; MUID:89313788
 A:Accession: B41682
 A:Molecule type: mRNA
 A:Residues: 3-607 <MOS>
 A:Cross-references: GB:M21442; NID:g213930; PIDN:AAA49637.1; PID:g213931
 R:Schorpp, M.; Doebbeling, U.; Wagner, U.; Ryffel, G.U.
 J. Mol. Biol. 199, 83-93, 1988
 A:Title: 5'-flanking and 5'-proximal exon regions of the two Xenopus albumin genes. Dele
 A:Reference number: S02692; MUID:88172470
 A:Accession: S02693
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-48 <SCH>

A:Cross-references: EMBL:Z26826
 R:Wolfe, A.P.; Glover, J.F.; Martin, S.C.; Tenniswood, M.P.R.; Williams, J.L.; Tata,
 Eur. J. Biochem. 146, 489-496, 1985
 A:Title: Deinduction of transcription of xenopus 74-kDa albumin genes and destabiliza
 A:Reference number: A05288; MUID:85126974
 A:Accession: A05288
 A:Molecule type: mRNA
 A:Residues: 459-502, 'L', 504-557 <WOL>
 A:Cross-references: GB:M28276
 A:Note: the authors translated the codon TAT for residue 63 as Thr
 C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds coppe
 mones (weak bonds with these hormones promote their transfer across the membranes), I
 C:Genetics: 27/1
 A:Introns: 27/1
 C:Superfamily: serum albumin; serum albumin repeat homology
 C:Keywords: carrier protein; duplication; glycoprotein; metal binding; plasma
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-24/Domain: propeptide #status predicted <PRO>
 F:25-607/Product: 74K serum albumin #status predicted <MAT>
 F:32-201/Domain: serum albumin repeat homology <SA1>
 F:220-393/Domain: serum albumin repeat homology <SA2>
 F:412-591/Domain: serum albumin repeat homology <SA3>
 F:30/Binding site: copper (His) #status predicted
 F:80-88, 101-117, 116-127, 147-192, 191-200, 223-269, 288-276, 288-302, 301-312, 339-384, 383-3
 F:256/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.9%; Score 77.5; DB 1; Length 607;
 Best Local Similarity 24.8%; Pred. No. 3.8;
 Matches 35; Conservative 17; Mismatches 48; Indels 41; Gaps 5;
 QY 4 DIVDQLKNVNDLVPEF-----LPAPDEVETNCSEAFSCF---QKALKSAN 48
 Db 72 EINDFAKSCINDTPECEKPVGTFFDKLCADPAVGNYEWSKCCAKQDPERAQCFKAH 131
 QY 49 TGNNERII---NVSIIKLRK-----PPSTNAGRRQKRLTQCP 84
 Db 132 RDHEHTSIKPEPEETCKLKEHPDDLLSAFIEARNHPDLYPPAVLALTKQYHKLAEHC 191
 QY 85 CDSYKPKPKPEFLERFKSLQ 105
 Db 192 CEEDKE--KCFSEKMKQLMK 210
 RESULT 4
 B71600
 variant-specific surface protein 1 homolog PFB1055C - malaria parasite (Plasmodium fa
 N:Alternate names: erythrocyte membrane binding protein 1 (EMPI)
 C:Species: Plasmodium falciparum
 C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
 C:Accession: B71600
 R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.
 ; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H
 Science 282, 1126-1132, 1998
 A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
 A:Reference number: A71600; MUID:99021743
 A:Accession: B71600
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-2197 <GAR>
 A:Cross-references: GB:AE001434; GB:AE001362; NID:g3845341; PIDN:AACT71996.1; PID:g384
 A:Experimental source: clone 3D7
 C:Genetics:
 A:Gene: PFB1055C

Query Match 13.3%; Score 74; DB 2; Length 2197;
 Best Local Similarity 29.5%; Pred. No. 32;
 Matches 28; Conservative 16; Mismatches 41; Indels 10; Gaps 4;
 QY 9 LKNVNDLVPEFLPAPDEVETNCSEAF--SCFQKALKSANTGNNGRIINVSIIKLRK 66
 Db 1594 VKSFLETWPK-IAVVDQDNVTKLSFGSCGSAISAISTN-GNEEDAIKMKLEKK 1651

QY 67 -----PPSTNAGRRQKHRLTCPCSDSYEKAPPKKE 95
Db 1652 IDECKRKFGGSGQTCTNLTHPLDVOEDEPLEE 1686

RESULT 5
T31527
hypothetical protein Y47D3A.14 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T31527
R:Matthews, L.
submitted to the EMBL Data Library, October 1999
A:Reference number: Z21043
A:Accession: T31527
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-868 <WIL>
A:Cross-references: EMBL:AL117202; PIDN:CAB55073.1; CESP:Y47D3A.14
A:Experimental source: clone Y47D3A
C:Genetics:
A:Gene: CESP:Y47D3A.14
A:Introns: 17/2; 51/3; 102/3; 169/3; 246/1; 287/2; 319/3; 371/3; 395/2; 440/3; 450/3; 440/3

Query Match 13.2%; Score 73.5; DB 2; Length 868;
Best Local Similarity 29.2%; Pred. No. 14;
Matches 19; Conservative 11; Mismatches 30; Indels 5; Gaps 1;

QY 33 WSAFSCFOKAOLKS-----ANTGNNERIINYSIKLKRPSTNAGRQKHRLTCPCSDS 87
||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 680 WSRVQAFRRKNNKMSGGGTSNSDSSEQSEKVLKKLRPSDETIVPPIHVICPLFPS 739

QY 88 YEKKP 92
: |
Db 740 HVAIP 744

RESULT 6
A32108
translation initiation factor eIF-2 alpha chain - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein J1429; protein YJR007w
C:Species: Saccharomyces cerevisiae
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 26-Aug-1999
C:Accession: A32108; S55195; S57022
R:Cigan, A.M.; Papich, E.K.; Feng, L.; Donahue, T.F.
Proc. Natl. Acad. Sci. U.S.A. 86, 2784-2788, 1989
A>Title: Yeast translation initiation suppressor su12 encodes the alpha-subunit of eukaryotic
A:Reference number: A32108; MUID:89202411
A:Accession: A32108
A:Molecule type: DNA
A:Residues: 1-304 <CTG>
A:Cross-references: EMBL:M25552; NID:g3411369; PIDN:AAA70332.1; PID:g903889
R:de Haan, M.; Smits, P.H.M.; Grivell, L.A.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55183
A:Accession: S55195
A:Molecule type: DNA
A:Residues: 1-304 <DEH>
A:Cross-references: EMBL:X87611; NID:g854567; PIDN:CAA60929.1; PID:g854580
R:de Haan, M.; Grivell, L.A.; Smits, P.H.M.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56771
A:Accession: S57022
A:Molecule type: DNA
A:Residues: 1-304 <ZAG>
A:Cross-references: EMBL:Z49507; NID:g1015631; PIDN:CAA89529.1; PID:g1015632; MIPS:YJR007w
C:Genetics:
A:Gene: SGD:SUIT2
A:Cross-references: SGD:S0003767; MIPS:YJR007w
A:Map position: 10R
C:Superfamily: translation initiation factor eIF-2 alpha chain

A;Residues: 1-262 <AVR>
A;Cross-references: GB:L22858; NID:g9510708; PIDN:AAA66699.1; PID:g559138
C;Genetics:
A;Gene: ACorf-69

Query Match 12.6%; Score 70; DB 2; Length 262;
Best Local Similarity 22.2%; Pred.No. 9.4;
Matches 18; Conservative 19; Mismatches 34; Indels 10; Gaps 2;

QY 30 NCWSAFSCFOAKLKSANTGNERRIINVSIKKLRKPPSTNAGRRQH-----RLTCPS 84
|| | : :: | -::| : ||| :
Db 176 NCVLKVDFADFEHETIQMLN----KFVNHFKWLYRPSPSRPANSERYLICFNKLVRPY 230
|| | : :: | -::| : ||| :

QY 85 CDSYEKKPKPEFLERFSLAQ 105
|-:-| :-| : -:| :
Db 231 CNIVNELEKOFKYRYRIQLK 251
|-:-| :-| : -:| :

RESULT 11
T41813
ACMNPV orf69 - Bombyx mori nuclear polyhedrosis virus (isolate T3)
C;Species: Bombyx mori nuclear polyhedrosis virus, BmSNPV
A;Variety: isolate T3
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C;Accession: T41813
R;Goni, S.; Majima, K.; Maeda, S.
J. Gen. Virol. 80, 1323-1337, 1999
A;Title: Sequence analysis of the genome of Bombyx mori nucleopolydovirus.
A;Reference number: Z22020; MUID:99281911
A;Accession: T41813
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A;Residues: 1-262 <KAM>
A;Cross-references: EMBL:L33180; PIDN:AAC63742.1
A;Experimental source: isolate T3
C;Genetics:
A;Note: Orf_57

Query Match 12.6%; Score 70; DB 2; Length 262;
Best Local Similarity 22.2%; Pred.No. 9.4;
Matches 18; Conservative 18; Mismatches 35; Indels 10; Gaps 2;

QY 30 NCWSAFSCFOAKLKSANTGNERRIINVSIKKLRKPPSTNAGRRQH-----RLTCPS 84
|| | : :: | -::| : ||| :
Db 176 NCVLKVDFADFEHKTIOMLN----KFVNHFKWLYRPSPSRPANSERYLICFNKLVRPY 230
|| | : :: | -::| : ||| :

QY 85 CDSYEKKPKPEFLERFSLAQ 105
|-:-| :-| : -:| :
Db 231 CNIVNELEKOFKYRYRIQLK 251
|-:-| :-| : -:| :

RESULT 12
I49124
interleukin-15 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 15-Mar-1996 #sequence_revision 15-Mar-1996 #text_change 16-Jul-1999
C;Accession: I49124
R;Anderson, D.M.; Johnson, L.; Glaccum, M.B.; Copeland, N.G.; Gilbert, D.J.; Jenkins
Genomics 25, 701-706, 1995
A;Title: Chromosomal assignment and genomic structure of IL15.
A;Reference number: A56005; MUID:95278940
A;Accession: I49124
A>Status: preliminary
A:Molecule type: mRNA
A;Residues: 1-162 <RES>
A;Cross-references: EMBL:U14332; NID:g984941; PIDN:AAAT5377.1; PID:g984942
C;Genetics:
A;Gene: IL15
C;Superfamily: interleukin-15

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	97.5	17.5	162	1	IL15_FELCA	O97687	felis silve
2	95.5	17.1	162	1	IL15_HUMAN	P40933	homo sapien
3	90.5	16.2	162	1	IL15_CERAE	P40221	cercopithec
4	90.5	16.2	162	1	IL15_MACMU	P48092	macaca mula
5	86.5	15.5	162	1	IL15_BOVIN	Q28028	bos taurus
6	82	14.7	337	1	IRL1_MOUSE	P14719	mus musculu
7	77.5	13.9	607	1	ALB2_XENLA	P14872	xenopus lae
8	76.5	13.7	162	1	IL15_PIG	Q95253	sus scrofa
9	72	12.9	304	1	IP2A_YEAST	P25459	saccharomyc
10	71.5	12.8	195	1	ALLA_YEAST	P32459	saccharomyc
11	70	12.6	162	1	Y069_NPVAC	P41469	autographa
12	69.5	12.5	162	1	IL15_MOUSE	P48346	mus musculu
13	69	12.4	206	1	YIM7_YEAST	P40470	saccharomyc
14	69	12.4	741	1	IDH_AZOVI	P16100	azotobacter
15	68	12.2	201	1	SDC2_HUMAN	P34741	homo sapien
16	67.5	12.1	300	1	SPY4_MOUSE	Q9wtp2	mus musculu
17	66.5	11.9	420	1	YIT3_DROME	P06607	drosophila
18	66.5	11.9	1750	1	Y832_METYA	Q25842	methanococc
19	66.5	11.9	3712	1	ACVS_CEPAC	P52464	cephalospor
20	66	11.8	805	1	AHR_MOUSE	P30561	mus musculu
21	65.5	11.8	576	1	P80C_HUMAN	P38432	homo sapien
22	65	11.7	695	1	TKT_PICST	P34736	pichia stip
23	64.5	11.6	162	1	IL15_RAT	P97604	rattus norv
24	64.5	11.6	484	1	CEB3_MOUSE	P54285	mus musculu
25	64.5	11.6	4385	1	YP73_CAEEL	Q09222	caenorhabdi
26	64	11.5	921	1	SVL_BACSU	Q45477	bacillus su
27	63.5	11.4	173	1	YNN3_YEAST	P53913	saccharomyc
28	63.5	11.4	457	1	ORD7_CAEEL	P41933	caenorhabdi
29	63.5	11.4	477	1	CEB3_RABIT	P54286	oryctolagus
30	63.5	11.4	484	1	CEB3_HUMAN	P54287	homo sapien
31	63.5	11.4	484	1	CEB3_RAT	Q03184	epstein-bar
32	63.5	11.4	525	1	UL32_EBV	P25558	saccharomyc
33	63.5	11.4	1636	1	BUD3_YEAST		


```

FT DISULFID          90   136      POTENTIAL.
FT CARBOHYD        127   127      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT           5       K -> T.
FT VARIANT         31     31     I -> T.
SQ SEQUENCE        162 AA; 18194 MW; D233CF7FF6188C01 CRC64;

Query Match                16.2%; Score 90.5; DB 1; Length 162;
Best Local Similarity    25.9%; Pred. No. 0.024;
Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;

Qy 3 IDIVDLKKNYNDLVPEF-----LPAPEDVETNCWMSAFSCF----OKAQJKSANTGNE 53
    :::: || | ::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 51 VNVISDLKK-TEDIQTSHRIDATLYTSDVHPSCVKVTAMCKFLLEQLQVISHESGDTHD 109

Qy 54 RIINVSIIKLKRKPSTNAGRQRHRLTCPSCDSYEKKPKPFERLERSLLQ 105
    : | : | | | : | | : | | : | | | | | | | | | | | | | | | | | | |
Db 110 TVENLI--LANNILSSNGNTES---GCKECELEEKNIKEFLQSFPVHIYO 156

RESULT 5
IL15_BOVIN STANDARD; PRT; 162 AA.
ID IL15_BOVIN AC Q28028;
DT DT 01-NOV-1997 (Rel. 35, Created)
DT DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN GN IL15.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HOLSTEIN;
RX MEDLINE=97426124; PubMed=9282828;
RA Canals A., Gasbarre L.C., Boyd P.C., Almeria S., Zarlenga D.S.;
RT "Cloning and expression of bovine interleukin-15: analysis and
modulation of transcription by exogenous stimulation.";
RL Interferon cytokine Res. 17:473-480(1997).
CC CC FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA (BY SIMILARITY).
CC -- SUBCELLULAR LOCATION: SECRETED.
CC -- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/ or
send an email to license@isb-sib.ch).
-----
EMBL; U42433; AAA85130.1; ..
Cytochrome; Glycoprotein; Signal.
DR SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48 POTENTIAL.
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 104 104
FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 162 AA; 18554 MW; 6A63CAA329EB8302 CRC64;

Query Match                15.5%; Score 86.5; DB 1; Length 162;
Best Local Similarity    25.5%; Pred. No. 0.061;

```



```

Query Match      13.7%; Score 76.5; DB 1; Length 162;
Best Local Similarity 24.1%; Pred. No. 0.61;
Matches 28; Conservative 19; Mismatches 42; Indels 27; Gaps 6;

QY  5 IVDOLKVVNDLVPEF-----LPAPEDVETNCESAFSCF-----QKAQLKSANTG 50
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  53 VISDLKK-IEDLIRSIHMADLYTESDAHPNCKVTAMKFCFLLELRVILQESRNSDISDTV 111
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY  51 NNERII-NVSIKKLKRPPSPNAGRRQKHRLTCSCDSYEEKPKPEFLERFKSLIQ 105
      | : | : | : : : | : | : | : | : | : | : | : | : | : | : | : |
Db  112 ENLIILANSLSISIEYK---TESG-----CKECEELEBKNI NEFLKSIHIHQ 156
      | : | : | : : : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 9
IF2A_YEAST
ID IF2A_YEAST STANDARD; PRT; 304 AA.
AC P20459;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE EUKARYOTIC TRANSLATION INITIATION FACTOR 2 ALPHA SUBUNIT (EIF-2-
  ALPHA).
GN SUI2 OR TIF211 OR YJR007W OR J1429.
GS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

```


FT	DISULFID	90	136	POTENTIAL.
FT	CARBOHYD	104	104	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	108	108	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	119	119	N-LINKED (GLCNAC. . .) (POTENTIAL).
QO	SEQUENCE	162 AA; 18593 MW; 58C971498CEBF296 CRC64;		

Query Match	12.5%;	Score 69.5;	DB 1;	Length 162;
Best Local Similarity	19.6%;	Pred. No. 3;		
Matches	21;	Conservative	29;	Mismatches 52; Indels 5; Gaps 3;

QY	3	IDI---	VDQKKNYNDL-VPFELPAPEDEVFNCEWSAFSCFQAQLKSANTGNNERIINV	58
	:	:	:	:
Db	51	IDRVYLEKIESLIQSHIDTTLVTSDFHPSCKVTAMNCF-LLEQLVILHEYSNMTLNE	109	

QY	59	SIKLKRKPPSTNAGRQKHRLTCPSDSYSEKKPKPEFLERFKSLQ	105	
	:	:	:	:
Db	110	TVRNVLANSTLSSNKNVAESGCKECEELEKTFTEFLQSFIRVQ	156	

RESULT	13			
YIM7_YEAST		STANDARD;	PRT;	206 AA.
ID	YIM7_YEAST			
AC	P40470;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	HYPOTHETICAL 23.8 KDA PROTEIN IN MET18-STH1 INTERGENIC REGION.			
GN	YIL127C.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.			
OX	NCBI_TaxID=4932;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN=S288C / AB972;			
RA	Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,			
RA	Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,			
RA	Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,			
RA	Louis E., Lyne G., Moule S., Moule T., Odell C., Pearson D.,			
RA	Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,			
RA	Walsh S.V., Whitehead S.;			
RL	Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; 246833; CAA86865.1; -			
DR	SGD; S0001389; YIL127C.			
KW	Hypothetical protein.			
QO	SEQUENCE	206 AA; 23845 MW; 4E51B11216F7ECB8 CRC64;		

Query Match	12.4%;	Score 69;	DB 1;	Length 206;
Best Local Similarity	26.0%;	Pred. No. 4.4;		
Matches	19;	Conservative	14;	Mismatches 40; Indels 0; Gaps 0;

QY	8	LKKNYNDLVPFELPAPEDEVFNCEWSAFSCFQAQLKSANTGNNERIINVSKLKRKP	67	
	:	:	:	:
Db	12	QATSVAGNLNLLPGVPKIRANNKTSVNGSKAQLIDRLNLRKRVQLQNRDVRHKKKC	71	

QY	68	PSTNAGRQKHRL	80	
	:	:	:	
Db	72	KLVRKKKKVKKHL	84	

RESULT	14			
IDH	AZOVIT			

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Lung fibroblast;
RX MEDLINE=89214123; PubMed=2523388;
RA Marynen P., Zhang J., Cassiman J.J., den Berghe H., David G.;
RT "Partial primary structure of the 48- and 90-kilodalton core proteins
of cell surface-associated heparan sulfate proteoglycans of lung
fibroblasts. Prediction of an integral membrane domain and evidence
for multiple distinct core proteins at the cell surface of human lung
fibroblasts."
RL J. Biol. Chem. 264:7017-7024(1989).
RC
CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -|- SIMILARITY: BELONGS TO THE SYNDSCAN FAMILY OF INTEGRAL MEMBRANE
CC PROTEOGLYCANS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
of the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; J04621; AAA5701.1; ALT_INIT.
CC
CC MIM; 142460;
CC InterPro; IPR001050;
CC Pfam; PF01034; Syndscan; 1.
CC PROSITE; PS00964; SYNDSCAN; 1.
KW Proteoglycan; Heparan sulfate; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 201 SYNDSCAN-2.
FT FT DOMAIN 19 144 EXTRACELLULAR (POTENTIAL).
FT FT TRANSMEM 145 169 POTENTIAL.
FT FT DOMAIN 170 201 CYTOPLASMIC (POTENTIAL).
FT SITE 142 143 CLEAVAGE OF ECTODOMAIN (POTENTIAL).
FT CARBOHYD 41 41 O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).
FT CARBOHYD 55 55 O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).
FT CARBOHYD 57 57 O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).
SQ SEQUENCE 201 AA; 22174 MW; 7B7F175650641054 CRC64;

Query Match 12.2%; Score 68; DB 1; Length 201;
Best Local Similarity 24.5%; Pred. No. 5.4;
Matches 24; Conservative 19; Mismatches 41; Indels 14; Gaps 2;

QY 7 DOLKNYNDLVPEFLPAPEDVETNCNSAFSCFOAKLKSANTGNE-----RIINV 58
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 108 DREKVHUSDSERKMDPAEEDINVTEKHSDSLFRTTEVLAAVIAGGVIGFLFAFLILL 167
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |

QY 59 SIKLKRKPPSTNAGRQRKRLTGPCSDSYSEKPKPEF 96
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 168 VYMRKRKDEGSDYDLGERK-----PSSAAAYQKAPTEF 199
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |

Search completed: May 23, 2001, 11:20:24
Job time: 522 sec

```

Search completed: May 23, 2001, 11:20:24
Job time: 522 sec

Result No.	Score	Query Match	Length	DB ID	Description
1	95.5	17.1	114	4	Q9UBA3
2	95.5	17.1	136	4	Q00440
3	85.5	15.4	162	6	Q9XSJ6
4	82	14.7	367	11	Q05208
5	77.5	13.9	3052	14	Q82933
6	75	13.5	566	5	Q9VVE8
7	74	13.3	2197	5	Q96296
8	73.5	13.2	868	5	Q9NAH8
9	71.5	12.8	187	13	Q9W756
10	71	12.7	1101	5	Q62471
11	70.5	12.7	848	4	Q75141
12	70	12.6	262	14	Q92434
13	69	12.4	1220	10	Q24015
14	68.5	12.3	336	11	Q62612
15	68.5	12.3	566	11	Q62811
16	68.5	12.3	1257	10	Q64316
17	68.5	12.3	2873	14	Q93069
18	68	12.2	329	1	Q28411
19	67.5	12.1	300	11	Q9WTP2
					Q9UBA3 homo sapien
					Q00440 homo sapien
					Q9XSJ6 ovnis aries
					Q05208 mus musculus
					Q82933 johnsongrass
					Q9VVE8 drosophila
					Q96296 plasmodium
					Q9NAH8 caenorhabdi
					Q9W756 gallus galli
					Q62471 caenorhabdi
					Q75141 homo sapien
					Q92434 bombyx mori
					Q24015 lycopersico
					Q62612 rattus norv
					Q62811 rattus norv
					Q64316 arabidopsis
					Q93069 hepatitis g
					Q28411 archaeoglob
					Q9WTP2 mus musculu

```

DT 01-JUL-1997 (TEMBLrel. 04, Last sequence update)
DT 01-AUG-1998 (TEMBLrel. 07, Last annotation update)
DE INTERLEUKIN-15.
GN IL-15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
OX [1]
RN SEQUENCE FROM N.A.
RP MEAZZA R., FERRINI S.;
RA Meazza R., Ferrini S.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y09908; CAA1044.1; -
SQ SEQUENCE 136 AA; 15259 MW; 494ED8DDC5D364CD CRC64;

Query Match 17.1%; Score 95.5; DB 4; Length 136;
Best Local Similarity 25.9%; Pred. No. 0.0072;
Matches 23; Conservative 22; Mismatches 46; Indels 15; Gaps 5;

QY 3 IDVQDKNYVNDLVPEF-----LPAPEDVETNCWSAFSCF----OKAQLKSANTGNE 53
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 25 VNVISDLKK-IEDLIQSMHIDATLYTSDVHPSCKVTKMCFLELQVLVISLESGDASIH 83
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 54 RIINVSIKKLRKPPSTNAGROKHRLTCTPCSDSYEKKKPKPEFLERKSLQ 105
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 84 TVENLIIT--LANNSLSSNGNWTES--GCKFCEBLEEKNIKEFLQSFVHIQ 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
Q9XSJ6 PRELIMINARY; PRT; 162 AA.
ID AC Q9XSJ6;
DT 01-NOV-1999 (TEMBLrel. 12, Created)
DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
DE INTERLEUKIN-15.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Casey G.J., Chaplin P.J.;
RT "Isolation of interleukin-15 mRNA transcripts from T and B cells
RT circulating in effluent lymph.";
RL J. Interferon Cytokine Res. 0:0-0(1999).
DR EMBL; AF149700; AAD37425.1; -
SQ SEQUENCE 162 AA; 18510 MW; 87195650E45F1E5D CRC64;

Query Match 15.4%; Score 85.5; DB 6; Length 162;
Best Local Similarity 28.8%; Pred. No. 0.1;
Matches 23; Conservative 10; Mismatches 46; Indels 1; Gaps 1;

QY 26 DVETNCWSAFSCFQAKLKSANTGNNERIINVSIIKKLRKPPSTNAGRRKHRLTGPCSC 85
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 78 DAHPNCKVTALQCFLELRLVLHESKNAAYEI-IENLTMADRLNLSIENKTELGCCEC 136
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 86 DSYEKRPKPEFLERKSLQ 105
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 137 EELEKSIKEFLKSFVHIQ 156
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
Q05Z08 PRELIMINARY; PRT; 567 AA.
ID AC Q05Z08;
DT 01-JUN-1998 (TEMBLrel. 06, Created)
DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)
DT 01-JUN-2000 (TEMBLrel. 14, Last annotation update)
DE ST2L PROTEIN PRECURSOR.

```

[illegible]

```
QY 67 -----PPSTNAGRQKHRLTCPCSDSYEKKPKPE 95
Db 1652 IDECRKPKGNSGTCNETLTHPLDVQDEPLEE 1686

RESULT 8
Q9NAH8 PRELIMINARY; PRT; 868 AA.
AC Q9NAH8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Y47D3A.14 PROTEIN.
GN Y47D3A.14
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Matthews L.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AL117202; CAB55073.1; -.
SQ SEQUENCE 868 AA; 100165 MW; 09A2A97066E373D7 CRC64;

Query Match 13.2%; Score 73.5; DB 5; Length 868;
Best Local Similarity 29.2%; Pred. No. 12;
Matches 19; Conservative 11; Mismatches 30; Indels 5; Gaps 1;

QY 33 WFAFCFOKAQLK-----ANTGNNERIINVSIIKLRKPPSTNAGRQKHRLTCPCSDS 87
Db 680 WSRVQAFRRKNNKMGSGGTSTSDSSEQSEKVLKLRKRRPSDETVPVPHVIVCPLPFS 739

QY 88 YEKKP 92
Db 740 HVAIP 744

RESULT 9
Q9W756 PRELIMINARY; PRT; 187 AA.
AC Q9W756;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (INTERLEUKIN 15 PRECURSOR).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SC; TISSUE=LIVER;
RA Burnside J.; Sofer L.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SC; TISSUE=LIVER;
RA Choi K.D.; Lillehoj H.S.; Burnside J.;
RT "Gallus gallus mRNA for IL-15 precursor.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF152927; AAD38392.1; -.
DR EMBL; AF139097; AAF61446.1; -.
SQ SEQUENCE 187 AA; 21964 MW; 470601EBF8837095 CRC64;

Query Match 12.8%; Score 71.5; DB 13; Length 187;
Best Local Similarity 25.3%; Pred. No. 3.8;
Matches 24; Conservative 10; Mismatches 34; Indels 27; Gaps 4;

QY 25 EDVETNCENSAFSCF---QKAQLKASANTGNNERIINVSIIKLRKPPSTNAGRQKHRLT 81
Db 98 EDIE--CQEPVMRCFFLEMKVILHECD-----IKKSRKHQDVNRNIWKNGARFA 144

QY 82 -----CPSCDSYSEKKPKPEFLERFKSLQ 105
Db 145 TYQLNSTTAKCKECEYEYKKNFTFQSFVKVIQ 179

RESULT 10
O62471 PRELIMINARY; PRT; 1101 AA.
AC O62471;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE Y45FI0B.10 PROTEIN.
GN Y45FI0B.10
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R.; Ainscough R.; Anderson K.; Baynes C.; Berks M.;
RA Bonfield J.; Burton J.; Connell M.; Copsey T.; Cooper J.; Coulson A.;
RA Craxton M.; Dear S.; Du Z.; Durbin R.; Favello A.; Fulton L.;
RA Gardner A.; Green P.; Hawkins T.; Hillier L.; Jier M.; Johnston L.;
RA Jones M.; Kershaw J.; Kirsten J.; Laister N.; Latreille P.;
RA Lightning J.; Lloyd C.; McMurray A.; Mortimore B.; O'Callaghan M.;
RA Parsons J.; Percy C.; Rifken L.; Roopra A.; Saunders D.; Showkeen R.;
RA Smaldon N.; Smith A.; Sonhammer E.; Staden R.; Sulston J.;
RA Thierry-Mieg J.; Thomas K.; Vaudin M.; Vaughan K.; Waterston R.;
RA Watson A.; Weinstock L.; Wilkinson-Sproat J.; Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994);
DR EMBL; AL021487; CAA16357.1; -.
DR INTERPRO; IPR001680; -.
DR PFAM; PF00400; WD40; 9.
DR PRINTS; PRO0320; GPROTEINBRPT.
DR PROSITE; PS00678; WD_REPEATS; UNKNOWN_2.
SQ SEQUENCE 1101 AA; 121014 MW; 9C9C557F8A67D6B0 CRC64;

Query Match 12.7%; Score 71; DB 5; Length 1101;
Best Local Similarity 31.8%; Pred. No. 28;
Matches 21; Conservative 12; Mismatches 23; Indels 10; Gaps 3;

QY 40 QKAQLKASANTGNNERIINVSIIKLRKPPSTNAGRQKHRLTCPCSDSYEKKPKPE 96
Db 1013 RSARAOSSVSSASNEPVASTSAGEIKDKPILSSNNGNAQSAPRATAP-----KPTFDM 1065

QY 97 LERFKS 102
Db 1066 LERSKS 1071

RESULT 11
O75141 PRELIMINARY; PRT; 848 AA.
ID O75141
AC O75141;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
```

```

QY 30 NCSEWSAFSCFQKAQLKSANTGNRRRIINYSIKLKRKPPSTNAGRQKH-----RLTCPS 84
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 176 NCVLKVDFAEHKTQIMLN-----KFNHFEXWLVKPPSSRRPANSERYLICFNKLVRPY 230
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 85 CDSYEKKPKPKFLERFKSLQ 105
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 231 CNDYVNELEKQFKYYRIQLK 251
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
O24015 PRELIMINARY; PRT: 1220 AA.
AC O24015;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE RESISTANCE COMPLEX PROTEIN I2C-1.
GS I2C-1.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97290204; PubMed=9144960;
RA Ori N., Eshed Y., Paran I., Presting G., Aviv D., Tanksley S.,
RA Zamir D., Fluhr R.;
RA "The I2C family from the wilt disease resistance locus I2 belongs to
RT the nucleotide binding, leucine-rich repeat superfamily of plant
RL resistance genes.";
RL Plant Cell 9:521-532(1997).
DR EMBL: AF004878; AAB63274.1; -
DR MENDEL; 24472; Lyces; 3172; 24472.
DR INTERPRO: IPR000767; -
DR INTERPRO: IPR001611; -
DR INTERPRO: IPR002182; -
DR PFAM: PF00560; LRR; 4.
DR PFAM: PF00931; NB-ARC; 1.
DR PRINTS: PR00364; DISEASERSIST.
SQ SEQUENCE 1220 AA; 139386 MW; B5997280C8C5EB20 CRC64;

Query Match 12.4%; Score 69; DB 10; Length 1220;
Best Local Similarity 30.3%; Pred. No. 51;
Matches 30; Conservative 8; Mismatches 23; Indels 38; Gaps

QY 13 VNDLVPEFLPA-----PEDVET-----NCSEWSAFSCFQKAQLKSANTGN 51
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 948 VDDISPEFLPTARQLSIENCHNVTRFLPTATSLHRCNEKLSMACGGAAQLTSLNIW- 1006
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 52 NERIINYSIKLKRKP-----PSTNAGRQKHRLT-CPSCD 86
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1007 -----GCKKLKCLPELLPSL-----KELRLTYCPEIE 1033
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
Q62612 PRELIMINARY; PRT: 336 AA.
AC Q62612
ID Q62612;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE FIT-1S PRECURSOR.
GS FIT-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94178260; PubMed=8131748;

```


Search completed: May 23, 2001, 11:19:42
Job time: 595 sec

```

Query Match      12.3%; Score 68.5; DB 11; Length 336;
Best Local Similarity 22.4%; Pred. No. 15;
Matches 28; Conservative 25; Mismatches 35; Indels 37; Gaps

QY      3 IDIVDLKKNVNDLVP-----EFLPAPEDVETNCWAFSCFQAQLKSN 48
      | : : : | : | : | : | : | : | : | : | : | : | : | : |
Db      49 INPVEWYVSYNTNERIPTQKRNRIFVSRDLKFLPAKVE-----DSGIYTCVIRSP-ESIK 102

QY      49 TGNNEELIINWSIKLKRKP-----STNAGRQRKHRLTCPCSDSYEKKPPKEFLE 98
      || : || : | : | : | : | : | : | : | : | : | : | : | : |
Db     103 TGS----LNVITYK---RPPNCKIPDMMYSYVDGSDKNSKTKTPTIALYNNTAPOWFK 155

```

QY	99 RFKSL 103 I:I
Db	156 NCKAL 160
RESULT 15	
Q62611	
ID	Q62611 PRELIMINARY; PRT; 566 AA.
AC	Q62611;
DT	01-NOV-1996 (TREMBLrel. 01, Created)
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE	FIT-IM PRECURSOR.
GN	FIT-1.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
NCBI	TaxID=10116;

[illegible]

```

Query Match      12.3%; Score 68.5; DB 11; Length 566;
Best Local Similarity 22.4%; Pred.No.25;
Matches 28; Conservative 25; Mismatches 35; Indels 37; Gaps

QY 3 IDIVDLKKNYVNDLVP-----EELPAPEDVETNCESAFSCFOKAQLKSAN 48
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
db 49 INPVENYVNTNRIETQKRNIFVSRDLKFLUPAKVE-----DSGIYTCVTRSP-ESIK 102

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:11:34 ; Search time 108.07 Seconds

(without alignments)
60.300 Million cell updates/sec

Title: US-09-522-217-2_COPY_32_145

Perfect score: 605

Sequence: 1 QDRMIRMRLIDIVDLKN.....DSYEKKPKPEFLERFKSLQ 114

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0401.*

1:	/SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2:	/SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3:	/SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4:	/SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5:	/SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6:	/SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7:	/SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8:	/SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9:	/SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10:	/SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11:	/SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12:	/SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13:	/SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14:	/SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15:	/SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16:	/SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17:	/SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18:	/SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19:	/SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20:	/SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21:	/SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22:	/SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	605	100.0	162	21	B18623 A human zalphall 1
2	605	100.0	519	21	B18627 Amino acid sequenc
3	351	58.0	146	21	B18624 A mouse zalphall 1
4	351	58.0	510	21	B18628 Amino acid sequenc
5	212	35.0	40	21	B18625 Antigeninc peptide
6	133	22.0	32	21	B18626 Antigeninc peptide
7	99.5	16.4	135	21	Y54825 Human Interleukin-
8	95.5	15.8	114	16	R83310 Human interleukin-
9	95.5	15.8	114	16	R83435 Mammalian interleu
10	95.5	15.8	114	16	R66928 Mammalian IL-15.
11	95.5	15.8	114	17	W09099 Simian mature epit

12	95.5	15.8	114	17	W07253 Generic mammalian
13	95.5	15.8	114	19	W39186 Simian epithelium
14	95.5	15.8	114	20	W39186 Simian epithelium
15	95.5	15.8	114	21	Y82309 Mature simian epit
16	95.5	15.8	162	16	R83438 Human interleukin-
17	95.5	15.8	162	16	R66927 Human IL-15. Homo
18	95.5	15.8	162	17	W09098 Simian epithelium
19	95.5	15.8	162	17	W07255 Human epithelium-d
20	95.5	15.8	162	17	R98527 Human interleukin-
21	95.5	15.8	162	18	W37369 Wild-type interleu
22	95.5	15.8	162	18	W01658 Human interleukin-
23	95.5	15.8	162	19	W53878 Human interleukin-
24	95.5	15.8	162	19	W39185 Simian epithelium
25	95.5	15.8	162	20	Y03756 Simian epithelium-
26	95.5	15.8	162	21	B18632 Amino acid sequenc
27	95.5	15.8	162	21	Y78595 Human interleukin-
28	95.5	15.8	162	21	Y52308 Simian epithelium-
29	95.5	15.8	162	22	B50870 Human IL-15. Homo
30	90.5	15.0	114	16	R83309 Simian interleukin
31	90.5	15.0	114	17	W09101 Human mature epith
32	90.5	15.0	114	19	W39188 Human epithelium d
33	90.5	15.0	114	20	Y03759 Human epithelium-d
34	90.5	15.0	114	21	Y52311 Mature human epith
35	90.5	15.0	122	17	R90842 Recombinant flag s
36	90.5	15.0	162	16	R83436 Simian interleukin
37	90.5	15.0	162	16	R66926 Simian IL-15. Cer
38	90.5	15.0	162	17	W09100 Human epithelium d
39	90.5	15.0	162	17	W07254 Simian epithelium-
40	90.5	15.0	162	17	R98526 Simian interleukin
41	90.5	15.0	162	17	R92798 Mammalian interleu
42	90.5	15.0	162	18	W37370 Mutant interleukin
43	90.5	15.0	162	19	W39187 Human epithelium d
44	90.5	15.0	162	20	Y03758 Human epithelium-d
45	90.5	15.0	162	21	Y78594 Simian interleukin

ALIGNMENTS

RESULT 1

B18623
ID B18623 standard; Protein; 162 AA.
XX
AC B18623;
XX
XX
DT 22-JAN-2001 (first entry)
XX
XX A human zalphall ligand polypeptide.
DE
XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
KW tumourigenesis; leukaemia; hematopoesis; B cell tumour.
XX
XX Homo sapiens.
OS
XX
XX WO200053761-A2.
XX
XX 14-SEP-2000.
XX
XX 09-MAR-2000; 2000WO-US06067.
XX
XX 09-MAR-1999; 99US-0264908.
PR 11-MAR-1999; 99US-0265992.
PR 01-JUL-1999; 99US-0142013.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Gross JA;
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX
XX WPI: 2000-565600/52.
DR N-PSDB: A75952.
XX
XX New human cytokine, designated zalphall ligand, useful for stimulating

PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX
 PS Disclosure; Page 205-206; 256pp; English.
 XX
 CC The present sequence represents a human zalphall ligand polypeptide,
 CC which is a cytokine. The zalphall ligand is useful for stimulating the
 CC proliferation and development of haematopoietic cells in vitro and in
 CC vivo. Zalphall ligand polynucleotides can be used as primers or probes
 CC for cloning the zalphall gene. The zalphall ligand is useful for
 CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
 CC used for treating leukaemias and lymphomas. Antagonists against zalphall
 CC ligand are useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 SQ Sequence 162 AA;

Query Match 100.0%; Score 605; DB 21; Length 162;
 Best Local Similarity 100.0%; Pred. No. 1.1e-63;
 Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QDRHMRQLDIDVDLKNVNDLVPEFLPAPEDVETNCWSEAFSCFQKALKSANTGN 60
 DB 32 qdrhmrmqlidivdqlknyvndlvpeflpapeuvtncwseafscfqkqlksantgn 91
 QY 61 NERIINVSIKLKRKPPSTNAGRQKHLTCPCSDSYEKKPKPEFLERFKSLQ 114
 DB 92 neriinvsikkrlkrppstnagrrqkhlrtcpscdsyekpkpkeflerfkslq 145

RESULT 2

B18627
 ID B18627 standard; Protein; 519 AA.
 XX
 AC B18627;
 XX
 DT 22-JAN-2001 (first entry)
 XX
 DE Amino acid sequence of MBP-human zalphall ligand fusion in pTAP126.
 XX
 KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopolesis; B cell tumour.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO200053761-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06067.
 XX
 PR 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 XX
 DR WPI; 2000-565600/52.
 DR N-PSDB; A75599.
 XX
 PT New human cytokine, designated zalphall ligand, useful for stimulating

PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX
 PS Example 31; Page 233-235; 256pp; English.
 XX
 CC The present sequence represents a MBP-human zalphall ligand fusion in
 CC the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
 CC is useful for stimulating the proliferation and development of
 CC haematopoietic cells in vitro and in vivo. Zalphall ligand
 CC polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 SQ Sequence 519 AA;

Query Match 100.0%; Score 605; DB 21; Length 519;
 Best Local Similarity 100.0%; Pred. No. 4.7e-63;
 Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QDRHMRQLDIDVDLKNVNDLVPEFLPAPEDVETNCWSEAFSCFQKALKSANTGN 60
 DB 389 qdrhmrmqlidivdqlknyvndlvpeflpapedvetncwseafscfqkqlksantgn 448
 QY 61 NERIINVSIKLKRKPPSTNAGRQKHLTCPCSDSYEKKPKPEFLERFKSLQ 114
 DB 449 neriinvsikkrlkrppstnagrrqkhlrtcpscdsyekpkpkeflerfkslq 502

RESULT 3

B18624
 ID B18624 standard; Protein; 146 AA.
 XX
 AC B18624;
 XX
 DT 22-JAN-2001 (first entry)
 XX
 DE A mouse zalphall ligand polypeptide.
 XX
 KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopolesis; B cell tumour.
 XX
 OS Mus musculus.
 OS
 PN WO200053761-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06067.
 XX
 PR 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 XX
 DR WPI; 2000-565600/52.
 DR N-PSDB; A75580.
 XX
 PT New human cytokine, designated zalphall ligand, useful for stimulating

PT	the proliferation and/or development of haematopoietic cells in vitro
PT	and in vivo, and for treating tumourigenesis -
XX	
PS	Disclosure; Page 222-223; 256pp; English.
XX	
CC	The present sequence represents a mouse zalphall ligand polypeptide,
CC	which is a cytokine. The zalphall ligand is useful for stimulating the
CC	proliferation and development of haematopoietic cells in vitro and in
CC	vivo. Zalphall ligand polynucleotides can be used as primers or probes
CC	for cloning the zalphall gene. The zalphall ligand is useful for
CC	treating tumourigenesis. A zalphall ligand-separin fusion toxin may be
CC	used for treating leukaemias and lymphomas. Antagonists against zalphall
CC	ligand are useful as research reagents for characterizing ligand-receptor
CC	interaction. Antagonists are also useful for inhibiting expansion,
CC	proliferation, activation and differentiation of cells involved in
CC	regulating hematopoiesis. The zalphall ligand may also be used to
CC	stimulate an immune response against B cell tumour, a virus, a parasite
CC	or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
CC	agonists and antibodies are also useful for the detection, diagnosis,
CC	prevention, and treatment of diseases associated with a zalphall ligand
CC	genetic defect.
XX	
Sequence	146 AA;

	Query Match	58.0%	Score 351;	DB 21;	Length 146;
	Best Local Similarity	59.3%;	Pred. No.	9.5e-34;	
	Matches	67;	Conservative	19;	Mismatches 27; Indels 0; Gaps 0;
QY	2 DRHMIRQLIDIVDQLKYNVNDLVPFLPAPEDVETNCEWSARSCFOKAOLKSANTGNN	61			
Dd	26 drrllirrhldivagikilyendldpellsapqdvkghehaafacfqaklkpsnpgnn	85			
QY	62 ERIINVSIKKLKPSTNACRQKHRLTCSDCDSYEKKPKPELERFKSLIQ	114			
	: : : : :				
Dd	86 ktfldidvaglrllrpalrraggkghakcpcdsyekrtokfejlrlkwll	138			

RESULT	4	
B18628	B18628	standard; Protein; 510 AA.
XX	AC	B18628;
XX	DT	22-JAN-2001 (first entry)
XX	XX	Amino acid sequence of MBP-mouse zalphall ligand fusion in pTAP134.
DE	DE	zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
KW	KW	tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
XX	XX	Synthetic.
OS	OS	Mus musculus.
XX	XX	WO200053761-A2.
PN	XX	14-SEP-2000.
PD	XX	09-MAR-2000; 2000WO-US060657.
PF	XX	09-MAR-1999; 99US-0264908.
PR	PR	11-MAR-1999; 99US-0265992.
PR	PR	01-JUL-1999; 99US-0142013.
XX	XX	(ZYMO) ZYMOGENETICS INC.
PA	XX	Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI	PI	Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX	XX	WPI; 2000-565600/52.
DR	DR	N-PSDB; A75602.
XX	XX	New human cytokine designated zalphall ligand, useful for stimulating

the proliferation and/or development of haematopoietic cells in vitro and in vivo, and for treating tumourigenesis -

Example 31; Page 239-240; 256pp; English.

The present sequence represents a MBP-mouse zalphall ligand fusion in the plasmid pPR126. Zalphall ligand is a cytokine. The zalphall ligand is useful for stimulating the proliferation and development of haematopoietic cells in vitro and in vivo. Zalphall ligand polynucleotides can be used as primers or probes for cloning the zalphall gene. The zalphall ligand is useful for treating tumourigenesis in vivo. The zalphall ligand fusion toxin may be used for treating leukaemias and lymphomas. Antagonists against zalphall ligand are useful as research reagents for characterizing ligand-receptor interaction. Antagonists are also useful for inhibiting expansion, proliferation, activation and differentiation of cells involved in regulating hematopoiesis. The zalphall ligand may also be used to stimulate an immune response against B cell tumour, a virus, a parasite or a bacterium. The zalphall polypeptides, polynucleotides, antagonists, agonists and antibodies are also useful for the detection, diagnosis, prevention, and treatment of diseases associated with a zalphall ligand genetic defect.

Sequence 510 AA;

```

Query Match      58.0%; Score 351; DB 21; Length 510;
Best Local Similarity 59.3%; pred. No. 4.7e-33;
Matches 67; Conservative 19; Mismatches 27; Indels 0; Gaps 0;

Qy 2 DRHIRMRLQIDIVDQLKNYVNDLVPELPAPEDVETNCESWSCFQKAQLKSANTGN 61
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 390 drllirhlldiveqlkiyendldpellsapqdgkghcehaatcfaklkpsnpgnn 449

Qy 62 ERIINVSITKKLRKPPSTNAGRGKHRUTCSQCSYEKKPKPEFLERPKSLLQ 114
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 450 ktflidvaglrirrlpqrqgkqkhhktakcsdcsvkrtokfejlrlkllq 502

```

RESULT	5	
B18625		
ID	B18625 standard; Peptide; 40 AA.	
XX	XX	
AC	B18625;	
XX	XX	
DT	22-JAN-2001 (first entry)	
XX	XX	
DE	Antigeninc peptide derived from a human zalphall ligand polypeptide.	
XX	XX	
KW	zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;	
KW	tumourigenesis; leukaemia; hematopolesis; B cell tumour.	
XX	XX	
OS	Homo sapiens.	
XX	XX	
PN	WO2000053761-A2.	
XX	XX	
PD	14-SEP-2000.	
XX	XX	
PF	09-MAR-2000; 2000WO-US060607.	
XX	XX	
PR	09-MAR-1999; 99US-0264908.	
PR	11-MAR-1999; 99US-0265992.	
PR	01-JUL-1999; 99US-0142013.	
XX	XX	
PA	(ZYMO) ZYMOGENETICS INC.	
XX	XX	
PI	Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;	
PI	Johnston JV, Nelson AJ, Dillon SR, Hammond AK;	
XX	XX	
XX	WPI; 2000-565600/52.	
XX	XX	
PT	New human cytokine, designated zalphall ligand, useful for stimulating	
PT	the proliferation and/or development of haematopoietic cells in vitro	

PT and in vivo, and for treating tumourigenesis -
XX Example 34; Page 227; 256pp; English.
PS
CC The present sequence was used to raise antibodies, and is derived from
CC a human zalphall ligand polypeptide. zalphall ligand is a cytokine.
CC The zalphall ligand is useful for stimulating the proliferation and
CC development of haematopoietic cells in vitro and in vivo. zalphall
CC ligand polynucleotides can be used as primers or probes for cloning the
CC a zalphall ligand-saporin fusion toxin may be used for treating
CC leukaemias and lymphomas. Antagonists against zalphall ligand are
CC useful as research reagents for characterizing ligand-receptor
CC interaction. Antagonists are also useful for inhibiting expansion,
CC proliferation, activation and differentiation of cells involved in
CC regulating hematopoiesis. The zalphall ligand may also be used to
CC stimulate an immune response against B cell tumour, a virus, a parasite
CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
CC agonists and antibodies are also useful for the detection, diagnosis,
CC prevention, and treatment of diseases associated with a zalphall ligand
CC genetic defect.
XX
SQ Sequence 40 AA;

Query Match 35.0%; Score 212; DB 21; Length 40;
Best Local Similarity 100.0%; Pred. NO. 4.8e-18;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDRHMRQLDIDVQLKNYVNDLVPFLPAPEDVETNC 40
DB 1 qdrhmrmqldivdqlknyvndlvpeflpapedvetc 40

RESULT 6
BI8626
ID BI8626 standard; Peptide; 32 AA.
XX
AC BI8626;
XX
XX 22-JAN-2001 (first entry)
XX
DE Antigeninc peptide derived from a human zalphall ligand polypeptide.
XX
XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
XX
XX Homo sapiens.
XX
XX W0200053761-A2.
XX
XX 14-SEP-2000.
XX
XX 09-MAR-2000; 2000WO-US06067.
XX
XX 09-MAR-1999; 99US-0264908.
PR 11-MAR-1999; 99US-0265992.
PR 01-JUL-1999; 99US-0142013.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX
XX WPI; 2000-565600/52.
XX
XX New human cytokine, designated zalphall ligand, useful for stimulating
PT the proliferation and/or development of haematopoietic cells in vitro
PT and in vivo, and for treating tumourigenesis -
XX
XX Example 34; Page 227; 256pp; English.
PS
XX The present sequence was used to raise antibodies, and is derived from

CC a human zalphall ligand polypeptide. zalphall ligand is a cytokine.
CC The zalphall ligand is useful for stimulating the proliferation and
CC development of haematopoietic cells in vitro and in vivo. zalphall
CC ligand polynucleotides can be used as primers or probes for cloning the
CC a zalphall ligand-saporin fusion toxin may be used for treating
CC leukaemias and lymphomas. Antagonists against zalphall ligand are
CC useful as research reagents for characterizing ligand-receptor
CC interaction. Antagonists are also useful for inhibiting expansion,
CC proliferation, activation and differentiation of cells involved in
CC regulating hematopoiesis. The zalphall ligand may also be used to
CC stimulate an immune response against B cell tumour, a virus, a parasite
CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
CC agonists and antibodies are also useful for the detection, diagnosis,
CC prevention, and treatment of diseases associated with a zalphall ligand
CC genetic defect.
XX
SQ Sequence 32 AA;

Query Match 22.0%; Score 133; DB 21; Length 32;
Best Local Similarity 100.0%; Pred. NO. 7.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 CPSCDSYEKKPKERFKSLQ 114
DB 1 cpscdsyekkpkkeflerfksllq 24

RESULT 7
Y54825
ID Y54825 standard; Protein; 135 AA.
XX
AC Y54825;
XX
XX 04-FEB-2000 (first entry)
XX
DE Human Interleukin-15 protein sequence.
XX
XX Interleukin-15; IL-15; antisense inhibitor; untranslated region; 5'UTR;
KW 3'UTR; human; infection; inflammation; tumour; therapy; diagnosis.
XX
XX Homo sapiens.
XX
XX US5985663-A.
PN
XX 16-NOV-1999.
PD
XX 25-NOV-1998; 98US-0200141.
PF
XX 25-NOV-1998; 98US-0200141.
PR
XX (ISIS-) ISIS PHARM INC.
PA
XX Bennett CF, Cowser LM;
PI
XX WPI; 2000-022283/02.
DR
XX N-PSDB; Z37358.
XX
XX Antisense compound useful for inhibiting human interleukin-15
PT expression useful for treating diseases associated with interleukin-15
PT expression -
XX
XX Example 13; Column 43-44; 31pp; English.
PS
XX This sequence is the human interleukin-15. The invention relates to
CC antisense compounds that are targeted to a 5' or 3' untranslated region
CC (5'UTR or 3'UTR) of a nucleic acid molecule encoding human interleukin-15
CC (IL-15), and inhibit the expression of human IL-15. The antisense
CC inhibitors are useful for inhibiting expression of IL-15 in human
CC cells or tissues in vitro, for treating humans or other animals suspected
CC of having or being prone to a disease associated with IL-15 expression,
CC e.g. infections, inflammation or tumours. The inhibitors can also be used

CC for research or diagnostic purposes. Using antisense compounds
 CC specifically and effectively inhibits IL-15 function.

XX Sequence 135 AA;

Query Match 16.4%; Score 99.5; DB 21; Length 135;
 Best Local Similarity 26.8%; Pred. No. 0.00044;
 Matches 30; Conservative 21; Mismatches 46; Indels 15; Gaps 5;
 QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF----QKAQLKSANTGNNE 62
 Db 24 vnvisdtkk-iedlqsmhidatlytesdvhpckvtamckflllelqvlsesgdasihd 82
 QY 63 RIINVSIIKKLRKPPSTNAGRQKHLTCPCSDSYEKPPKPEFLERFKSLIQ 114
 Db 83 tvenlii--lannslssngvntes---gkceeeleeknikelfqsfvhiq 129

RESULT 8
 ID R83310 standard; Protein; 114 AA.
 AC R83310;
 DT 02-FEB-1996 (first entry)
 DE Human interleukin-15 mature polypeptide.
 KW Interleukin-15; epithelium-derived T-cell factor; T lymphocyte.
 OS Homo sapiens.
 PN WO9527722-A.
 PD 19-OCT-1995.
 PF 06-APR-1994; 94WO-US03793.
 PR 06-APR-1994; 94WO-US03793.
 PA (IMMV) IMMUNEX CORP.
 PI Anderson DM, Eisenman JR, Fung V, Grabstein KH;
 PI Rauch C;
 DR WPI: 1995-373556/48.
 DR N-PSDB; T00527.

Isolated DNA encoding polypeptide with mammalian IL-15 activity - which stimulates proliferation and differentiation of T cells, used for treating carcinoma(s), melanomas, etc. and viral infections
 Claim 23; Page 30; 48pp; English.
 A simian species of IL-15 (sIL-15) was purified and its AA sequence and cDNA sequence analysed (see R83309, R83436, T00524, T00525). Both the simian and the human ORFs encode a precursor polypeptide (R83436, R83438). The precursor polypeptides each comprise a 48-AA leader sequence and a sequence encoding mature simian or human IL-15 polypeptides. The active simian and human IL-15 polypeptides are disclosed in R83309 & R83310 respectively. The invention also comprises other mammalian IL-15, including human IL-15, that hybridise to probes defined by R83438. A plasmid contg. a recombinant clone of human IL-15 cDNA was deposited with ATCC on Feb. 19 1993 under ATCC 69245. The deposit was named 141-hETF. R83435 is a mammalian mature IL-15 polypeptide. It is a generic sequence which encompasses both R83309 (simian) and R83310 (human) IL-15 mature polypeptides.
 Sequence 114 AA;

Query Match 15.8%; Score 95.5; DB 16; Length 114;
 Best Local Similarity 25.9%; Pred. No. 0.0011;
 Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;

QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF----QKAQLKSANTGNNE 62
 Db 3 vnvisdtkk-iedlqsmhidatlytesdvhpckvtamckflllelqvlsesgdasihd 61
 QY 63 RIINVSIIKKLRKPPSTNAGRQKHLTCPCSDSYEKPPKPEFLERFKSLIQ 114
 Db 62 tvenlii--lannslssngvntes---gkceeeleeknikelfqsfvhiq 108

RESULT 9
 ID R83435 standard; Protein; 114 AA.
 AC R83435;
 DT 02-FEB-1996 (first entry)
 DE Mammalian interleukin-15 mature polypeptide.
 KW Interleukin-15; epithelium-derived T-cell factor; T lymphocyte.
 OS Mammalian.
 FH Key Location/Qualifiers
 FT Misc-difference 52 /label= L,H
 FT Misc-difference 57 /label= A,T
 FT Misc-difference 58 /label= S,D
 FT Misc-difference 73 /label= S,I
 FT Misc-difference 80 /label= V,I

WO9527722-A.
 19-OCT-1995.
 06-APR-1994; 94WO-US03793.
 06-APR-1994; 94WO-US03793.
 (IMMV) IMMUNEX CORP.
 Anderson DM, Eisenman JR, Fung V, Grabstein KH;
 Rauch C;
 WPI: 1995-373556/48.

Isolated DNA encoding polypeptide with mammalian IL-15 activity - which stimulates proliferation and differentiation of T cells, used for treating carcinoma(s), melanomas, etc. and viral infections
 Claim 1; Page 33; 48pp; English.
 A simian species of IL-15 (sIL-15) was purified and its AA sequence and cDNA sequence analysed (see R83309, R83436, T00524, T00525). Both the simian and the human ORFs encode a precursor polypeptide (R83436, R83438). The precursor polypeptides each comprise a 48-AA leader sequence and a sequence encoding mature simian or human IL-15 polypeptides. The active simian and human IL-15 polypeptides are disclosed in R83309 & R83310 respectively. The invention also comprises other mammalian IL-15, including human IL-15, that hybridise to probes defined by R83438. A plasmid contg. a recombinant clone of human IL-15 cDNA was deposited with ATCC on Feb. 19 1993 under ATCC 69245. The deposit was named 141-hETF. R83435 is a mammalian mature IL-15 polypeptide. It is a generic sequence which encompasses both

CC R83309 (simian) and R83310 (human) IL-15 mature polypeptides.
 XX
 SQ Sequence 114 AA;

Query Match 15.8%; Score 95.5; DB 16; Length 114;
 Best Local Similarity 25.2%; Pred. No. 0.0011;
 Matches 29; Conservative 22; Mismatches 43; Indels 21; Gaps 5;
 Qy 12 IDVDQLKNYNDLVPEF-----LPAPEDVETNCWSAFSCF-OQAQLKSANTGNN---- 61
 Db 3 vnvisdlikk-iedliqsmhidatlytesdvhpckvtamkcflllelqvixesgdxixhd 61
 Qy 62 --ERIINVSIIKKLRKPPSTNAGRRQKRLTCPCSDSYEKPKPFLEFRFKSLIQ 114
 Db 62 tvenlilannxlssngnxtesg-----ckeceeleeknikelfsqfwhivq 108

RESULT 10
 ID R66928 standard; Protein; 114 AA.
 AC R66928;
 DT 04-SEP-1995 (first entry)
 DE Mammalian IL-15.
 KW Interleukin-15; IL-15; sIL-15; T-cell growth factor;
 KW antitumor; virucide.
 OS Mammalia.

Key Location/Qualifiers
 FT Misc-difference 52
 FT /label= Leu, His
 FT Misc-difference 57
 FT /label= Ala, Thr
 FT Misc-difference 58
 FT /label= Ser, Asp
 FT Misc-difference 73
 FT /label= Ser, Ile
 FT Misc-difference 80
 FT /label= Val, Ile

XX ZA9402636-A.
 XX
 XX 28-DEC-1994.
 XX 18-APR-1994; 94ZA-0002636.
 XX 18-APR-1994; 94ZA-0002636.
 XX (IMMV) IMMUNEX CORP.
 XX Anderson DM, Eisenman JR, Fung V, Grabstein KH;
 XX Rauch C;
 XX WPI; 1995-082473/11.
 XX
 XX New purified interleukin-15 - which induces T cell proliferation
 XX and differentiation, used for the treatment of tumours and viral
 XX infection
 XX
 XX Claim 1; Page 33; 47pp; English.
 XX
 XX Simian and human IL-15 cDNAs (Q84583-84) can be used to obtain
 CC cDNAs encoding other mammalian homologs of IL-15. A general
 CC sequence of mammalian IL-15 is claimed.
 XX
 SQ Sequence 114 AA;

Query Match 15.8%; Score 95.5; DB 16; Length 114;
 Best Local Similarity 25.2%; Pred. No. 0.0011;
 Matches 29; Conservative 22; Mismatches 43; Indels 21; Gaps 5;
 Qy 12 IDVDQLKNYNDLVPEF-----LPAPEDVETNCWSAFSCF-OQAQLKSANTGNN---- 61
 Db 3 vnvisdlikk-iedliqsmhidatlytesdvhpckvtamkcflllelqvixesgdxixhd 61
 Qy 62 --ERIINVSIIKKLRKPPSTNAGRRQKRLTCPCSDSYEKPKPFLEFRFKSLIQ 114
 Db 62 tvenlilannxlssngnxtesg-----ckeceeleeknikelfsqfwhivq 108

RESULT 11
 ID W09099 standard; Protein; 114 AA.
 AC W09099;

DT 11-MAR-1997 (first entry)
 DE Simian mature epithelium derived T cell factor.
 KW sETf; African green monkey; CV1/EBNA cell; T-cell; B-cell;
 KW lymphocyte; proliferation; differentiation; gastrointestinal;
 KW HIV infection; human immunodeficiency virus.
 OS Cercopithecus aethiops.

Key Location/Qualifiers
 FT Protein
 FT 1..114
 FT /label= mature_sETf

XX US5574138-A.

XX 12-NOV-1996.

XX 08-MAR-1993; 93US-0031399.

XX 22-FEB-1995; 95US-0393305.

XX 08-MAR-1993; 93US-0031399.

XX 22-APR-1994; 94US-0233606.

XX (IMMV) IMMUNEX CORP.

XX Anderson DM, Eisenman JR, Fung V, Grabstein KH;
 XX Rauch C;

XX WPI; 1996-517923/51.

XX N-PSDB; T49455.

XX New epithelium derived T cell factor - induces proliferation of T
 XX cells and stimulates destruction of tumour and virus-infected
 XX cells and protects against toxicity, partic. for treating intestinal
 XX disease and HIV infection

XX Claim 1; Fig 1; 35pp; English.

XX The simian ETF (epithelium derived T cell factor) was isolated from
 CC African green monkey CV1/EBNA cell conditioned medium. The N-
 CC terminal sequence of the purified sETf was determined and then PCR
 CC primers were designed based on the sequence information. A 92 bp
 CC fragment was amplified from CV1/EBNA DNA and was used as a probe to
 CC screen a CV1/EBNA cDNA library for the full-length sETf coding
 CC sequence. Mature sETf induces proliferation and/or differentiation
 CC of precursor or mature T cells and is useful for promoting long-term
 CC in vitro culture of T-lymphocytes and T-cell lines. It is used for
 CC treating gastrointestinal diseases including peptic ulcer, colitis and
 CC malignancy and for treating HIV infection.

XX Sequence 114 AA;

Best Local Similarity	25.28;	Pred. No. 0.0011;	
Matches	29;	Conservative	22; Mismatches 43; Indels 21; Gaps 5;
Qy	12	IDIYDLKNIYNDLVPEP-----LPAPEDVETNCWSAFSCF-QKAOIKSANTGN	--- 61
Db	3	vnvisdlkk-iedlqsmhidatlytesdvhpsscvtamckfllelqlvisesgdxixhd	61
Qy	62	--ERIINVSIIKKRKPPSTNAGRQRHRLTQPCSDSYKPKPPKFEFLERKSLQ	114
Db	62	tvenlilannxlssngnxtesg-----ckeceeleeknikelfqsfvhiq	108
RESULT	13		
ID	W39186		
XX	W39186	standard; Protein; 114 AA.	
AC	W39186;		
XX	08-MAY-1998	(first entry)	
XX	Simian epithelium derived T-cell factor mature protein.		
XX	Epithelium derived T-cell factor; ETF; simian; gastrointestinal disease;		
KW	B-cell stimulation; epithelium; villi; CD4+; CD8+; T-cell stimulation;		
KW	treatment; prevention.		
XX	Simian.		
XX	Key	Location/Qualifiers	
FT	Protein	1..114	
FT		/label= ETF	
XX	US5707616-A.		
XX	13-JAN-1998.		
XX	04-OCT-1996;	96US-0726817.	
XX	22-FEB-1995;	95US-0393305.	
PR	08-MAR-1993;	93US-0031399.	
PR	22-APR-1994;	94US-0233606.	
XX	(IMMV) IMMUNEX CORP.		
PA	Anderson DM, Eisenman JR, Fung V, Grabstein KH;		
PI	Rauch C;		
XX	WPI; 1998-100295/09.		
DR	N-PSDB; V02873.		
DR	Treatment or prevention of gastrointestinal diseases - by		
PT	administering epithelium-derived T-cell factor polypeptide		
XX	Claim 1a; Column 37-38; 34pp; English.		
XX	This sequence represents a simian epithelium-derived T-cell factor (ETF)		
CC	mature protein which is used in a method for treating or preventing		
CC	gastrointestinal disease. These polypeptides have particular application		
CC	in the treatment of gastrointestinal disorders associated with disruption		
CC	of the gastrointestinal epithelium or villi such as chemotherapy- and		
CC	radiation-therapy induced enteritis (gut toxicity), mucositis, peptic		
CC	ulcer disease, gastroenteritis and colitis, villus atrophic disorders,		
CC	malignancy and inflammatory bowel disease. ETF polypeptides may also be		
CC	useful in the treatment of human immunodeficiency virus (HIV) and		
CC	HIV-associated disease due to their ability to stimulate CD4+ and CD8+		
CC	cells. Biologically active ETF may be used to treat a variety of other		
CC	diseases or conditions where T-cell or B cell stimulation is desired.		
XX	Sequence	114 AA;	50

Query Match . 15.8%; Score 95.5; DB 19; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.0011;

Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;

Qy 12 IDIVDQLKKNYNDLVPEF-----LPAPEDVETNCWEAFSCF-----OKAQKLSANTGNNE 62
 Db 3 vnvisdllk-iedllqsmhidatlytesdvhpkskvtkamckflllelqvlslesgdasidh 61
 Qy 63 RIINVSIIKKLRKPPSTNAGRQRHRLTQPCSDSYEKPKPEFLERFKSLIQ 114
 Db 62 tvenlii--lannslsngnvtes---gckeceeleeknikelflqsfvhiq 108

RESULT 14
 Y03757
 ID Y03757 standard; Protein; 114 AA.
 AC Y03757;
 XX
 DT 10-JUN-1999 (first entry)
 XX
 DE Simian epithelium-derived T-cell factor (ETF) mature protein sequence.
 XX
 KW Epithelium-derived T-cell factor; ETF; immunoassay; enteritis; ulcer;
 KW T cell proliferation; gastrointestinal disease; mucositis; colitis;
 KW gastroenteritis; villus atrophy; malignancy; inflammatory bowel disease;
 KW human immune deficiency virus; tumour; simian.
 XX
 OS Mammalia.
 PN US5892001-A.
 XX
 XX 06-APR-1999.
 XX
 XX 04-OCT-1996; 96US-0725969.
 XX
 XX 22-FEB-1995; 95US-0393305.
 PR 08-MAR-1993; 93US-0031399.
 PR 22-APR-1994; 94US-0233606.
 PR 04-OCT-1996; 96US-0725969.
 XX
 PA (IMMV) IMMUNEX CORP.
 XX
 XX Anderson DM, Eisenman JR, Fung V, Grabstein KH;
 PI Rauch C;
 XX
 XX WPI; 1999-253930/21.
 DR N-PSDB; X29479.
 XX
 XX Antibodies specific for epithelium-derived T-cell growth factor
 PT
 XX
 PS Claim 1; Fig 1; 34pp; English.
 XX
 CC The invention relates to an isolated antibody that binds specifically to
 CC a simian or human epithelium-derived T-cell factor (ETF) polypeptide. The
 CC antibodies are used, optionally when immobilized or labeled, to detect
 CC and quantify ETF in standard immunoassays. They may also be used as
 CC diagnostic and therapeutic agents, e.g. when conjugated to toxins (or
 CC their precursors) or radionuclides. ETF induces proliferation and/or
 CC differentiation of T cells (or their precursors), e.g. for use in
 CC establishing long term in vitro cultures; and is also used to treat
 CC gastrointestinal disease (e.g. enteritis or mucositis induced by
 CC chemotherapy or radiation, peptide ulcer, gastroenteritis, colitis,
 CC villus atrophy, malignancy and inflammatory bowel disease), to treat
 CC human immune deficiency virus infection or associated disease, or
 CC generally in any situation requiring stimulation of T or B cell
 CC proliferation, secretion of immunoglobulins or certain cytokines,
 CC increased anti-infectious disease immunity, induction of T-cell lytic
 CC activity or increased destruction of tumour or virus-infected cells. The
 CC present sequence represents the mature active sequence of simian ETF
 CC polypeptide.
 XX
 XX Sequence 114 AA;
 SO

Query Match 15.8%; Score 95.5; DB 20; Length 114;
 Best Local Similarity 25.9%; Pred. No. 0.0011;
 Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;

Qy 12 IDIVDQLKKNYNDLVPEF-----LPAPEDVETNCWEAFSCF-----OKAQKLSANTGNNE 62
 Db 3 vnvisdllk-iedllqsmhidatlytesdvhpkskvtkamckflllelqvlslesgdasidh 61
 Qy 63 RIINVSIIKKLRKPPSTNAGRQRHRLTQPCSDSYEKPKPEFLERFKSLIQ 114
 Db 62 tvenlii--lannslsngnvtes---gckeceeleeknikelflqsfvhiq 108

RESULT 15
 Y52309
 ID Y52309 standard; Protein; 114 AA.
 XX
 AC Y52309;
 XX
 DT 09-FEB-2000 (first entry)
 XX
 DE Mature simian epithelium-derived T-cell factor (ETF).
 XX
 KW ETF; epithelium-derived T-cell factor; T-cell; T-lymphocyte;
 KW proliferation; differentiation; growth factor; precursor; mature; CD4+;
 KW CD8+; gastrointestinal; epithelium; cell culture; treatment; HIV;
 KW gastrointestinal disease; gastroenteritis; colitis;
 KW inflammatory bowel disease; villus atrophic disorder; enteritis;
 KW chemotherapy; radiotherapy; gut toxicity; cancer; side effect;
 KW tolerated dose.
 XX
 OS Mammalia.
 PN US5985262-A.
 XX
 XX 16-NOV-1999.
 XX
 XX 03-FEB-1997; 97US-0794524.
 XX
 XX 22-FEB-1995; 95US-0393305.
 PR 04-OCT-1996; 96US-0726817.
 PR 08-MAR-1993; 93US-0031399.
 PR 22-APR-1994; 94US-0233606.
 XX
 PA (IMMV) IMMUNEX CORP.
 XX
 XX Rauch C, Fung V, Eisenman JR, Grabstein KH, Anderson DM;
 PI WPI; 2000-022267/02.
 DR N-PSDB; 238244.
 XX
 XX Stimulation of T-cells in human immunodeficiency virus infected
 PT patients -
 XX
 PS Claim 1; Fig 1; 33pp; English.
 XX
 CC This sequence represents mature simian epithelium-derived T-cell factor
 CC (ETF). This is a previously unidentified T-cell growth factor which
 CC stimulates precursor or mature CD4+ and CD8+ T-lymphocytes to proliferate
 CC and differentiate. It also promotes proliferation of the gastrointestinal
 CC epithelium. The protein can be used to promote long-term in vitro culture
 CC of T-lymphocytes and T-cell lines. ETF can be used for treating HIV
 CC infection, HIV-associated diseases, and other diseases or conditions
 CC where stimulation of T-cell proliferation would be desirable e.g., it
 CC could be used to augment the destruction of tumour cells or virally-
 CC infected cells. ETF may also be used to treat or prevent gastrointestinal
 CC disease, including chemotherapy and radiotherapy associated enteritis,
 CC gastroenteritis, colitis, inflammatory bowel disease and villus atrophic
 CC disorders. Chemotherapy and radiotherapy associated enteritis (gut
 CC toxicity) results in bleeding and sepsis due to gastrointestinal flora
 CC entering the blood, and thus can limit the dosage of therapeutic agent
 CC administered to a cancer patient. ETF may therefore be used to increase
 CC the tolerated doses radiotherapy and chemotherapy.

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	95.5	15.8	114	1	US-08-031-399-6		Sequence 6, Appli
2	95.5	15.8	114	1	US-08-031-399-12		Sequence 12, Appl
3	95.5	15.8	114	1	US-08-393-305-3		Sequence 3, Appli
4	95.5	15.8	114	1	US-08-726-817-3		Sequence 3, Appli
5	95.5	15.8	114	1	US-08-726-817-6		Sequence 6, Appli
6	95.5	15.8	114	1	US-08-504-042-6		Sequence 6, Appli
7	95.5	15.8	114	1	US-08-504-042-12		Sequence 12, Appl
8	95.5	15.8	114	2	US-08-725-969-3		Sequence 3, Appli
9	95.5	15.8	114	4	US-08-794-524-3		Sequence 3, Appli
10	95.5	15.8	114	4	US-09-189-193-3		Sequence 3, Appli
11	95.5	15.8	114	5	PCN-US94-03793-6		Sequence 6, Appli
12	95.5	15.8	162	1	PCN-US94-03793-12		Sequence 12, Appl
13	95.5	15.8	162	1	US-08-031-399-5		Sequence 5, Appli
14	95.5	15.8	162	1	US-08-393-305-2		Sequence 2, Appli
15	95.5	15.8	162	1	US-08-535-733-2		Sequence 2, Appli
16	95.5	15.8	162	1	US-08-726-817-2		Sequence 2, Appli
17	95.5	15.8	162	1	US-08-504-042-5		Sequence 5, Appli
18	95.5	15.8	162	2	US-08-725-969-2		Sequence 2, Appli
19	95.5	15.8	162	3	US-08-842-947-6		Sequence 6, Appli
20	95.5	15.8	162	4	US-09-189-193-2		Sequence 2, Appli
21	95.5	15.8	162	5	PCN-US94-03793-5		Sequence 5, Appli
22	95.5	15.8	162	5	PCN-US96-06423-2		Sequence 2, Appli
23	90.5	15.0	114	1	US-08-031-399-3		Sequence 3, Appli
24	90.5	15.0	114	1	US-08-393-305-6		Sequence 6, Appli
25	90.5	15.0	114	1	US-08-726-817-6		Sequence 6, Appli
26	90.5	15.0	114	1	US-08-504-042-3		Sequence 3, Appli
27	90.5	15.0	114	2	US-08-725-969-6		Sequence 6, Appli

Db	3	VNVISDLKK - IEDLIQSMHIDATLYTESDVHPSCKVTAMKCFLELQVLSLESGDASIH	61
QY	63	RIINYSIKLKKPPSTWAGRQXHRUTCPSDSEYKPPKPELFERKSLQ	114
Db	62	TWENLII--LANNSSUNGNVTES---GCKECELEBKNIKEFIQSFVHIQ	108

```

RESULT      2
US-08-031-399-12
: Sequence 12, Application US/08031399
: Patent No. 552303
: GENERAL INFORMATION:
: APPLICANT: Grabstein, Kenneth
: APPLICANT: Anderson, Dirk
: APPLICANT: Eisenman, June
: APPLICANT: Fung, Victor
: APPLICANT: Rauch, Charles
: TITLE OF INVENTION: Epithelium-derived T-cell Factor
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Immunex Corporation
: STREET: 51 University Street
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98101

```

```

;
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/031,399
; FILING DATE: 19930308
; CLASSIFICATION: 530

```

Query Match	15.8%	Score	95.5	DB 1:	Length	114
Best Local Similarity	25.2%	Pred. NO.	0.0003			
Matches	29	Conservative	22	Mismatches	43	Gaps
				Indels	21	

Qy 12 IDIVDLKVVNDLVPEF-----LPAPDETNCWSAFSCF-QKAQLKSANTGN--- 61
::: || : ||: | || : | : | : | : | :
Db 3 VNVISDLKK-IEDLQTSMHIDATLYTESDVHPCKVTAMKGELLEQLVVISXESGDXXIHD 61

```

RESULT      3
US-08-393-305-3
: Sequence 3, Application US/083933005
: Patent No. 5574138
: GENERAL INFORMATION:
: APPLICANT: Grabstein, Kenneth
: APPLICANT: Anderson, Dirk
: APPLICANT: Eisenman, June
: APPLICANT: Fung, Victor

```

```

: APPLICANT: Rauch, Charles
: TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Seed and Berry
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98104
:

```

```

:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
:

```

APPLICATION NUMBER: US/08/393,305
FILING DATE: 22-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:

ATTORNEY/AGENT INFORMATION:
NAME: McMasters, David D.
REGISTRATION NUMBER: 33,9
REFERENCE/DOCKET NUMBER:

Query Match 15.8%; Score 95.5; DB 1; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.0003;
Matches 29; Conservative 22; Mismatches 46; Indels 1

QY 12 IDIVDLKNVNDLVPEF'-----LPAPEDVTNCEWSAFSCF-----QKAQLKSANT
::: || : || : | | : : | | : | : | : | : | : | :
Db 3 VNVTSDLKK -IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELLOVLVSLESGDA

RESULT 4
US-08-726-817-3
; Sequence 3, Application US/08726817
; Patent No. 5707616

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

OY	63 RIINVSIIKKLRKPPSTNAGRQKHRLTCPSCDSYEKKPPEFLERFKSLQ	114
DH	62 MVNIIT - FANNICISCNCHNTES - --CCKRCEET DEPVNTIPEOCSEVHTU	109

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-504-042-12

Query Match          15.8%; Score 95.5; DB 1; Length 114;
Best Local Similarity 25.2%; Pred. No. 0.0003;
Matches 29; Conservative 22; Mismatches 43; Indels 2

QY      12 IDIVDLKNYVDLVPEF-----LPAPEDVETNCWSAFSCF-QKAQLKSANTGN
       :::: ||| ::||| |   ||| :|: ||| :|: ||| :|: ||| :
```

Query Match 15.8%; Score 95.5; DB 1; Length 114;
Best Local Similarity 25.2%; Pred. No. 0.0003;
Matches 29; Conservative 22; Mismatches 43; Indels 2

QY 12 IDVDLKNYVNDLVEF-----LPAPEDVETNCWSAFSCF-QKAQKLSKANTGN


```
;;
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/189,193
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/393,305
;; FILING DATE: 22-FEB-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Mcmasters, David D.
;; REGISTRATION NUMBER: 33,963
;; REFERENCE/DOCKET NUMBER: 480052.409C2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-622-4900
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 114 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-09-189-193-3
```

```
Query Match 15.8%; Score 95.5; DB 4; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.0003;
Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;

QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVTNCWSAFSCF-----QKAQLKSANTGNNE 62
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLEQLQVISLESQDASIH 61

QY 63 RIINVSIIKKLRKPPSTNAGRQKHRLTQPCSDSYEKKPPKPEFLERFKSLQ 114
Db 62 TVENLII--LANNSSNGNVTES---GCKECELEEKNIKEFLQSFVHIQ 108
```

```
RESULT 10
PCT-US94-03793-6
; Sequence 6, Application PC/TUS9403793
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Interleukin-15
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03793
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
```

```
;;
;; LENGTH: 114 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: NO
;; PCT-US94-03793-6

Query Match 15.8%; Score 95.5; DB 5; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.0003;
Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;

QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVTNCWSAFSCF-----QKAQLKSANTGNNE 62
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLEQLQVISLESQDASIH 61

QY 63 RIINVSIIKKLRKPPSTNAGRQKHRLTQPCSDSYEKKPPKPEFLERFKSLQ 114
Db 62 TVENLII--LANNSSNGNVTES---GCKECELEEKNIKEFLQSFVHIQ 108
```

```
RESULT 11
PCT-US94-03793-12
; Sequence 12, Application PC/TUS9403793
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Interleukin-15
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03793
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; PCT-US94-03793-12
```

```
Query Match 15.8%; Score 95.5; DB 5; Length 114;
Best Local Similarity 25.2%; Pred. No. 0.0003;
Matches 29; Conservative 22; Mismatches 43; Indels 21; Gaps 5;

QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVTNCWSAFSCF-QKAQLKSANTGNN---- 61
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLEQLQVISLESQDXXIH 61

QY 62 --ERIINVSIIKKLRKPPSTNAGRQKHRLTQPCSDSYEKKPPKPEFLERFKSLQ 114
```


Db 62 TVENLIILANNXLSNGNXGTSF-----CKECELEEKNIKEFLQSFVHIQ 108

RESULT 12

US-08-031-399-5
; Sequence 5, Application US/08031399
; Patent No. 5552303
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Epithelium-derived T-cell Factor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/031.399
; FILING DATE: 19930308
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-031-399-5

Query Match 15.8%; Score 95.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.00047;
Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;
Qy 12 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCESAFSCF-----KQAOLKSANTGNNE 62
Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTKAMKCFLEQLQVLSLESGDASIH 109
Qy 63 RIINVSIIKKLRKPPSTNAGRQKRLTCPCSDSYEKPKPFLEFKSLIQ 114
Db 110 TVENLIIL--LANNSSNGNVTES---GCKECELEEKNIKEFLQSFVHIQ 156

RESULT 13

US-08-393-305-2
; Sequence 2, Application US/08393305
; Patent No. 5574138
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393.305
; FILING DATE: 22-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-393-305-2

Query Match 15.8%; Score 95.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.00047;
Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;
Qy 12 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCESAFSCF-----KQAOLKSANTGNNE 62
Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTKAMKCFLEQLQVLSLESGDASIH 109
Qy 63 RIINVSIIKKLRKPPSTNAGRQKRLTCPCSDSYEKPKPFLEFKSLIQ 114
Db 110 TVENLIIL--LANNSSNGNVTES---GCKECELEEKNIKEFLQSFVHIQ 156

RESULT 14

US-08-535-733-2
; Sequence 2, Application US/08535733
; Patent No. 5660824
; GENERAL INFORMATION:
; APPLICANT: Quinn, LeBris
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Troutt, Anthony B.
; TITLE OF INVENTION: Muscle-Trophic Factor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: System 7, Word 5.1a
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/535.733
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Malaska, Stephen L.
; REGISTRATION NUMBER: 32,655
; REFERENCE/DOCKET NUMBER: 2833

S29498
 S29498
 lymphocyte antigen Ly84 precursor - mouse
 N:Alternate names: 38.5K T1 glycoprotein; ST2L protein
 C:Species: Mus musculus (house mouse)
 C:Date: 13-Jan-1995 #sequence-revision 13-Jan-1995 #text_change 01-Dec-2000
 C:Accession: S29498; A33541; S17657; S07054
 R:Fanigisawa, K.; Takagi, T.; Tsukamoto, T.; Tetsuka, T.; Tomihaga, S.
 FEBS Lett. 318, 83-87, 1993
 A:Title: Presence of a novel primary response gene ST2L, encoding a product
 A:Reference number: S29498; MUID:93170492
 A:Accession: S29498
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-567 <Y>
 A:Cross-references: ENBL;D13695; NID:q286100; PIDN:BAA02854.1; PID:q286101

QY 76 -----PPSTNAGRRQKHRLTCPCSDSYEKKPKPE 104
DB 1652 IDBCKRRKPGSGQTCNETLTHPLDVQDEPLEE 1686

RESULT 5
T31527
hypothetical protein Y47D3A.14 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T31527
R:Matthews, L.
submitted to the EMBL Data Library, October 1999
A:Reference number: Z21043
A:Accession: T31527
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-868 <MIL>
A:Cross-references: EMBL:AL117202; PIDN:CAB55073.1; CESP:Y47D3A.14
A:Experimental source: clone Y47D3A
C:Genetics:
A:Gene: CBSP:Y47D3A.14
A:Introns: 17/2; 51/3; 102/3; 169/3; 246/1; 287/2; 319/3; 371/3; 395/2; 440/3; 450/3; 49

Query Match 12.1%; Score 73.5; DB 2; Length 868;
Best Local Similarity 29.2%; Pred. No. 19;
Matches 19; Conservative 11; Mismatches 30; Indels 5; Gaps 1;

QY 42 WSAFSCFQAKLKS-----ANTGNERRIINVSKLKRKPPSTNAGRRQKHRLTCPCSDS 96
DB 680 WSRVQAFRRKNNKSGGGTSNSDSQSEKVLKLRKRPSTVLPVPHYIVCPPLFPS 739

QY 97 YEKKP 101
DB 740 HVAIP 744

RESULT 6
C69230
conserved hypothetical protein WTH973 - Methanobacterium thermoautotrophicum (strain Del
C:Species: Methanobacterium thermoautotrophicum
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Sep-1999
C:Accession: C69230
R:Smith, D.R.; Doucette-Stamm, L.A.; Delouchery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.;
ki, S.; Church, G.W.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funcu
A:Reference number: A69000; MUID:98037514
A:Accession: C69230
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-237 <MTH>
A:Cross-references: GB:AE000871; GB:AE000666; NID:g2622069; PIDN:AA85469.1; PID:g262207
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: WTH973
A:Start codon: GTG
C:Superfamily: hypothetical protein MJ0915

Query Match 11.9%; Score 72; DB 2; Length 257;
Best Local Similarity 25.0%; Pred. No. 7.8;
Matches 27; Conservative 12; Mismatches 29; Indels 40; Gaps 3;

QY 5 MIRMQLDIDVQKKNY-----VNDLVPPELPAPEDV----- 36
DB 48 MVDGRAVLVDADMPLVOLVVEAASPEAVRDVLPPEILEAGKDVVMSVGLMDPELREML 107

QY 37 -----ETNCENSAFSCFQAKLKSANTGNERRIINVSKLKRKPP 77
DB 1652 IDBCKRRKPGSGQTCNETLTHPLDVQDEPLEE 1686

DB 108 VELASLNDATIHVPSGAIVGLDGLKAASMGNTIE-----SVKLIIRKPP 150

RESULT 7
A32108
translation initiation factor eIF-2 alpha chain - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein J1429; protein YJR007w
C:Species: Saccharomyces cerevisiae
C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 26-Aug-1999
C:Accession: A32108; S55195; S57022
R:Cigan, A.M.; Pabich, E.K.; Feng, L.; Donahue, T.F.
Proc. Natl. Acad. Sci. U.S.A. 86, 2784-2788, 1989
A:Title: Yeast translation initiation suppressor su12 encodes the alpha-subunit of eu
A:Reference number: A32108; MUID:89202411
A:Accession: A32108
A:Molecule type: DNA
A:Residues: 1-304 <CIG>
A:Cross-references: EMBL:M25552; NID:g341369; PIDN:AAA70332.1; PID:g903889
R:de Haan, M.; Smits, P.H.M.; Grivell, L.A.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55183
A:Accession: S55195
A:Molecule type: DNA
A:Residues: 1-304 <DEH>
A:Cross-references: EMBL:X87611; NID:g854567; PIDN:CAA60929.1; PID:g854580
R:de Haan, M.; Grivell, L.A.; Smits, P.H.M.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56771
A:Accession: S57022
A:Molecule type: DNA
A:Residues: 1-304 <ZAG>
A:Cross-references: EMBL:Z49507; NID:g1015631; PIDN:CAA89529.1; PID:g1015632; MIPS:YJ

A:Gene: SGD:SUI2
A:Cross-references: SGD:S0003767; MIPS:YJR007w
A:Map position: 10R
C:Superfamily: translation initiation factor eIF-2 alpha chain
C:Keywords: phosphoprotein; protein biosynthesis

Query Match 11.9%; Score 72; DB 2; Length 304;
Best Local Similarity 33.3%; Pred. No. 9.3;
Matches 20; Conservative 14; Mismatches 22; Indels 4; Gaps 4;

QY 13 DIVDLQKNYVND-LVPELPAPEDVETNC-EMSAFSCFQAKLKSANTGNERRIINVSK 70
DB 164 DVLDLANVYSKRLTPQAVKIRADVENSCFSYEGIDAIKDA-LKSAEDMSTEQ-MQVKV 221

RESULT 8
S42022
ureidoglycolate hydrolase (EC 3.5.3.19) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YIR032c
C:Species: Saccharomyces cerevisiae
C>Date: 31-Mar-1992 #sequence_revision 14-Sep-1994 #text_change 21-Jul-2000
C:Accession: S42022; S48494
R:Yoo, H.S.; Cooper, T.G.
Yeast 7, 693-698, 1991
A:Title: The ureidoglycolate hydrolase (DAL3) gene in Saccharomyces cerevisiae.
A:Reference number: S42022; MUID:92133160
A:Accession: S42022
A:Molecule type: DNA
A:Residues: 1-195 <YOO>
A:Cross-references: EMBL:M64778; NID:g171369; PIDN:AAA73025.1; PID:g171370
R:Rowley, K.
submitted to the EMBL Data Library, October 1994
A:Reference number: S48478
A:Accession: S48494
A:Molecule type: DNA
A:Residues: 1-195 <ROW>
A:Cross-references: GB:Z47047; EMBL:238061; NID:g603997; PID:g763377; MIPS:YIR032c
C:Genetics:
A:Gene: SGD:DAL3

A:Cross-references: SGD:S0001471; MIPS:YIR032c

A:Map position: 9R
C:Keywords: allantoin degradation; hydrolase; lipoprotein; methylated carboxyl end; pred
F:192/Blinding site: farnesyl (Cys) (covalent) #status predicted
F:192/Modified site: methyl ester carboxyl end (Cys) (in mature form) #status predicted

Query Match 11.8%; Score 71.5; DB 2; Length 195;

Best Local Similarity 29.2%; Pred. No. 6.6;
Matches 19; Conservative 9; Mismatches 24; Indels 13; Gaps 1;

Qy 15 VDQKNYVNDLVPEFLPAPEDVETNCWESAFSCFQAKLKSANTGNRIINVSIIKKLR 74

Db 48 VSQVENKSTSKVP-----NWNLFRCFPQPHLNRVFTQGSNQALSHSIRVLEK 94

Qy 75 KPPST 79

Db 95 HPCST 99

RESULT 9

A83317 isocitrate dehydrogenase PA2624 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: A83317

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337

A:Accession: A83317

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-741 <SFO>

A:Cross-references: GB:AE004691; GB:AE004091; NID:99948688; PIDN:AAG06012.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: idh; PA2624

C:Superfamily: Vibrio isocitrate dehydrogenase (NADP+)

Query Match 11.8%; Score 71.5; DB 2; Length 741;

Best Local Similarity 27.9%; Pred. No. 26;

Matches 24; Conservative 17; Mismatches 32; Indels 13; Gaps 3;

Qy 2 DRHMRQLIDIVDQKNYVNDL-----VPEFLPAPEDVETNCWESAFSCFQAKLKS 55

Db 78 DANIKLPNISASVPLQGAETAEGLGYKVPDF---PEDPOTDEEKEVRARYAKILGSA 134

Qy 56 AN----TGNRIINVSIIKKLRKPP 77

Db 135 VNPVLREGNSDRAPAAVKARKHP 160

RESULT 10

T27977

lin-15A protein - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000

C:Accession: T27977; S47893

R:Kershaw, J.

submitted to the EMBL Data Library, August 1996

A:Reference number: Z20449

A:Accession: T27977

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-719 <WIL>

A:Cross-references: EMBL:Z79605; PIDN:CAB01904.1; GSPDB:GN000028; CESP:ZK678.1

A:Experimental source: clone ZK678

R:Clark, S.G.; Lu, X.; Horvitz, H.R.

Genetics 137, 987-997, 1994

A:Title: The Caenorhabditis elegans locus lin-15, a negative regulator of a tyrosine

A:Reference number: S47893; MUID:95073612

A:Accession: S47893

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-719 <CLA>

A:Cross-references: EMBL:U10411; NID:g516013; PIDN:AAA20087.1; PID:g516014

C:Genetics:

A:Gene: CESP:ZK678.1

A:Map position: X

A:Introns: 93/3; 123/3; 269/3; 304/3; 360/3

C:Superfamily: Caenorhabditis elegans lin-15A protein

Query Match 11.7%; Score 71; DB 2; Length 719;

Best Local Similarity 23.0%; Pred. No. 28;

Matches 31; Conservative 17; Mismatches 43; Indels 44; Gaps 5;

Qy 3 RHMRQLIDIVD-----QLKNYVNDLVPEFLPAPED----- 35

Db 255 RHE-RLRMCYDHVDFVYEMCMDAFEAKTESEINEMPPDRLMRGHDYRALKRVGDLHKG 313

Qy 36 VETNCWESAFSCFQAKLKSANTGNRIINVSIIK-----LKRKPPSTNAGR 83

Db 314 VTSNPLYSF-----KNSIKSYRNHVPRVNGSLSPKPFSELVALLOSVPSTLINE 369

Qy 84 RQKRLTCPCDSYE 98

Db 370 LLNHNLSLSDADKQE 384

RESULT 11

T26919

hypothetical protein Y45F10B.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T26919

R:McMurray, A.

submitted to the EMBL Data Library, January 1998

A:Reference number: Z20286

A:Accession: T26919

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1101 <WIL>

A:Cross-references: EMBL:AL021487; PIDN:CAA16357.1; GSPDB:GN000022; CESP:Y45F10B.10

A:Experimental source: clone Y45F10B

C:Genetics:

A:Gene: CESP:Y45F10B.10

A:Map position: 4

A:Introns: 217/3; 292/1; 451/3; 556/2; 637/3; 755/3; 823/2; 946/1; 1078/3

Query Match 11.7%; Score 71; DB 2; Length 1101;

Best Local Similarity 31.8%; Pred. No. 44;

Matches 21; Conservative 12; Mismatches 23; Indels 10; Gaps 3;

Qy 49 QKAQLKSANTGNRIINVSIIKKRKP--PSTNAGRRQK-HRLTCPCDSYEKKPKPEF 105

Db 1013 RSARAGSVSSASNEPVASTSAGEIKKDPILSSNNGGNAQSAPRATAP-----KPTFDM 1065

Qy 106 LERFKS 111

Db 1066 LERSKS 1071

RESULT 12

T00372

hypothetical protein KIAA0650 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 17-Nov-2000

C:Accession: T00372; F12523

R:ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura,
DNA Res. 5, 169-176, 1998

Result No.	Query Match	Score	Length	DB	ID	%		Description
						Match	Score	
1	97.5	16.1	162	1	IL15_FELCA	097687	felis silve	
2	95.5	15.8	162	1	IL15_HUMAN	P40933	homo sapien	
3	90.5	15.0	162	1	IL15_CERAE	P40221	cercopitheci	
4	90.5	15.0	162	1	IL15_MACMU	P48092	macaca mula	
5	86.5	14.3	162	1	IL15_BOVIN	Q28028	bos taurus	
6	82	13.6	337	1	IRL1_MOUSE	P14719	mus musculus	
7	78	12.9	741	1	IDH_AGOVI	P16100	azotobacter	
8	77.5	12.8	607	1	ALB2_XENLA	P14872	xenopus lae	
9	76.5	12.6	162	1	IL15_PIG	Q95253	mus scrofa	
10	72	11.9	304	1	IF2A_YEAST	P20459	saccharomyc	
11	71.5	11.8	195	1	ALIA_YEAST	P32459	saccharomyc	
12	70	11.6	262	1	Y069_NPVAC	P41469	autographa	
13	70	11.6	403	1	IFT3_MOUSE	Q64345	mus musculus	
14	69.5	11.5	162	1	IL15_MOUSE	P48346	mus musculus	
15	69	11.4	206	1	YIM7_YEAST	P40470	saccharomyc	
16	68.5	11.3	162	1	IL15_RAT	P97604	rattus norv	
17	68	11.2	201	1	SPC2_HUMAN	P34741	homo sapien	
18	67.5	11.2	300	1	SV4_MOUSE	Q947P2	mus musculus	
19	66.5	11.0	420	1	VIT3_DROME	P06607	drosophila	
20	66.5	11.0	1636	1	BUD3_YEAST	P25558	saccharomyc	
21	66.5	11.0	1750	1	Y832_METYA	Q58242	methanococc	
22	66.5	11.0	3712	1	ACVS_CEPAC	P25464	cephalospor	
23	66	10.9	805	1	AHR_MOUSE	P30561	mus musculus	
24	65.5	10.8	576	1	P80C_HUMAN	P38432	homo sapien	
25	65	10.7	695	1	TKT_PICST	P34736	pichia strip	
26	64.5	10.7	484	1	CCB3_MOUSE	P54285	mus musculus	
27	64.5	10.7	1206	1	FM14_MOUSE	Q05859	mus musculus	
28	64.5	10.7	4385	1	YF73_CAEEL	Q09222	caenorhabdi	
29	64	10.6	921	1	SYI_BACSU	Q45477	bacillus su	
30	63.5	10.5	173	1	YNN3_YEAST	P53913	saccharomyc	
31	63.5	10.5	441	1	FUS6_ARATH	P45432	arabidopsis	
32	63.5	10.5	457	1	ODR7_CAEEL	P41933	caenorhabdi	
33	63.5	10.5	472	1	IFT2_HUMAN	P09913	homo sapien	

```

-----
or send an email to license@isb-sib.ch).
-----
CC CC
DR EMBL; U14407; AAA21551.1;
DR EMBL; X91233; CAAG2616.1;
DR EMBL; X94223; CAAG3914.1;
DR EMBL; X94222; CAAG3913.1;
DR EMBL; AF031167; AAB97518.1;
DR MIM; 600554;
KW Cytokine; Glycoprotein; Signal; Alternative splicing.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48 POTENTIAL.
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 127 127 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARSPLIC 1 37 MRISKPLRSISITQCYCLLLNSHFLTEAGIHVFILG ->
FT CONFLICT 141 141 MVLGTIDLCS (IN ISOFORM IL15-S21AA).
FT SEQUENCE 162 AA; 18086 MW; 0CE0520C1D8379E2 CRC64;
SQ
Query Match 15.8%; Score 95.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.011;
Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;
QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKALKSANTGNNE 62
Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTAMKCFLELQVLSLESGDASHD 109
QY 63 RIINVSIKLKRKPPSTNAGRRKOKHRLTCPCSDSYKKPKPKFLERFKSLIQ 114
Db 110 TVENLII--LANNLSNGNVTES---GCKECELEEKNIKEFLQSFVHIQV 156
RESULT 3
IL15_CERAE STANDARD; PRT; 162 AA.
ID IL15_CERAE
AC P40221;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 49-81.
RC TISSUE=Kidney;
RX MEDLINE=94233380; PubMed=8178155;
RA Grabstein K.H., Eisenman J., Shanebeck K., Rauch C.,
RA Srinivasan S., Fung V., Beers C., Richardson J., Schoenborn M.A.,
RA Andieh M., Johnson L., Alderson M.R., Watson J.D., Anderson D.M.,
RA Giri J.G.;
RT "Cloning of a T cell growth factor that interacts with the beta chain
of the interleukin-2 receptor.";
RL Science 264:965-968(1994).
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----

```

```
CC -----
DR EMBL; U03099; AA18416.1;
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48
FT CHAIN 49 162 INTERLEUKIN-15.
FT FT DISULFID 83 133 POTENTIAL.
FT FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 127 137 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 162 AA; 18222 MW; 1BF9A82644E1C9B7 CRC64;

Query Match 15.0%; Score 90.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.034;
Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;

QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCESAFSCF-----QKAOLKSANTGNNE 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCKVTAMKCFLELQVISHESGDTDIHD 109
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 RIINVSIKKLRKPPSTNAGRQKHRLTCPCSDSYEKKPKPEFLERFKSLQ 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 110 TVENLII--LANNILSSNGNITES---GCKECELEEKNIKEFLQSFVHIVQ 156
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
IL15_MACMU
ID IL15_MACMU STANDARD; PRT; 162 AA.
AC P48092;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=96003435; PubMed=7561102;
RA Villinger F.J., Brar S.S., Mayne A.E., Chikkala N., Ansari A.A.;
RT "Comparative sequence analysis of cytokine genes from human and
RT nonhuman primates.";
RL J. Immunol. 155:3946-3954(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Tatsumi M.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U19843; AAB60398.1;
DR EMBL; AAB00555; BAA19149.1;
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48
FT CHAIN 49 162 INTERLEUKIN-15.
FT FT DISULFID 83 133
SQ SEQUENCE 162 AA; 18554 MW; 6A63CAA329EB8302 CRC64;
```

```
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 127 127 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARIANT 5 5 K -> T.
FT VARIANT 31 31 I -> T.
SQ SEQUENCE 162 AA; 18194 MW; D233CF7FF6188C01 CRC64;

Query Match 15.0%; Score 90.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.034;
Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;

QY 12 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCESAFSCF-----QKAOLKSANTGNNE 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCKVTAMKCFLELQVISHESGDTDIHD 109
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 RIINVSIKKLRKPPSTNAGRQKHRLTCPCSDSYEKKPKPEFLERFKSLQ 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 110 TVENLII--LANNILSSNGNITES---GCKECELEEKNIKEFLQSFVHIVQ 156
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
IL15_BOVIN
ID IL15_BOVIN STANDARD; PRT; 162 AA.
AC Q28028;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HOLSTEIN;
RX MEDLINE=97426124; PubMed=9282828;
RA Canals A., Gasbarre L.C., Boyd P.C., Almeria S., Zarlenaga D.S.;
RT "Cloning and expression of bovine interleukin-15: analysis and
RT modulation of transcription by exogenous stimulation.";
RL J. Interferon Cytokine Res. 17:473-480(1997).
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U42433; AAA85130.1;
DR EMBL; U42433; AAA85130.1;
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48
FT CHAIN 49 162 INTERLEUKIN-15.
FT FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 104 104 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 113 113 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 121 121 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 127 127 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 162 AA; 18554 MW; 6A63CAA329EB8302 CRC64;

Query Match 14.3%; Score 86.5; DB 1; Length 162;
Best Local Similarity 25.5%; Pred. No. 0.084;
```

Matches 27; Conservative 16; Mismatches 56; Indels 7; Gaps 3;

QY 14 IVDQKNVNDLVPEF-----LPAPEDVETNCWSAFSCFOKAQLKSANTGNRIINVS 68
 DB 53 VINDLKT-IEHLQSIHMDATVYTESDAHNPCKVTAMQCFLLELRLVILHESKNATYIEI- 110
 QY 69 IKLKKRPPSTNAGRRKHRLTCPCSDSYEKPKPEFLERFKSLQ 114
 DB 111 IENLTMLANSLSIENKTGCKCEBELEBSKEFLUKSFVHIQV 156

RESULT 6
 ID IRL1_MOUSE STANDARD; PRT; 337 AA.
 AC P14719;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE INTERLEUKIN 1 RECEPTOR-LIKE 1 PRECURSOR (ST2 PROTEIN) (TI PROTEIN)
 DE (LYMPHOCYTE ANTIGEN 84).
 GN IL1RL1 OR ST2 OR ST2 OR LY84.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C;
 RX MEDLINE=90092495; PubMed=2532153;
 RA Tomimaga S.;
 RT "A putative protein of a growth specific cDNA from BALB/c-3T3 cells
 RT is highly similar to the extracellular portion of mouse interleukin 1
 RT receptor.";
 RL FEBS Lett. 258:301-304(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C3H/HE; TISSUE=Spleen;
 RX MEDLINE=91355215; PubMed=1832015;
 RA Tomimaga S.I.; Jenkins N.A.; Gilbert D.J.; Copeland N.G.;
 RA Tetsuka T.;
 RT "Molecular cloning of the murine ST2 gene. Characterization and
 RT chromosomal mapping.";
 RL Biochim. Biophys. Acta 1090:1-8(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89345536; PubMed=2527364;
 RA Klemenz R.; Hoffmann S.; Werenskiold A.K.;
 RT "Serum- and oncoprotein-mediated induction of a gene with sequence
 RT similarity to the gene encoding carcinoembryonic antigen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:5708-5712(1989).
 CC -!- FUNCTION: POSSIBLY INVOLVED IN REGULATION OF T-LYMPHOCYTE
 CC ACTION.
 CC -!- DEVELOPMENTAL STAGE: GROWTH-SPECIFIC EXPRESSION, G1-PHASE OF
 CC CELL CYCLE.
 CC -!- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -!- SIMILARITY: STRONG, TO INTERLEUKIN-1 RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: Y07519; CAA68812.1;
 CC EMBL: X60184; CAA42742.1;
 CC EMBL: M24843; AAA40160.1;
 CC PIR: S07054; S07054.
 CC MGD: MGI:98427; Ly84.
 CC InterPro: IPR003006;
 CC Pfam: PF00047; Ig; 3.
 KW Immunoglobulin domain; Glycoprotein; Signal.

FT SIGNAL 1 26
 FT CHAIN 27 337
 FT DOMAIN 35 100
 FT DOMAIN 132 194
 FT DOMAIN 233 315
 FT DISULFID 42 93
 FT DISULFID 139 187
 FT DISULFID 240 308
 FT CARBOHYD 60 60
 FT CARBOHYD 101 101
 FT CARBOHYD 107 107
 FT CARBOHYD 146 146
 FT CARBOHYD 176 176
 FT CARBOHYD 194 194
 FT CARBOHYD 225 225
 FT CARBOHYD 259 259
 FT CARBOHYD 278 278
 FT VARIANT 192 192
 SQ SEQUENCE 337 AA; 38502 MW; 757437272486926 CRC64;

Query Match 13.6%; Score 82; DB 1; Length 337;
 Best Local Similarity 27.1%; Pred. No. 0.52;
 Matches 29; Conservative 17; Mismatches 29; Indels 32; Gaps 6;

QY 16 DQKNVNDLVPEFLPAPEDVETNCWSAFSCFOKAQLKSANTGNRIINVSIKLKRK 75
 DB 76 DRLK-----FLPARVE-----DSGIYACV-----IRSPNL-NKTGYLNVTHIK---K 113
 QY 76 PPSSTN-----AGRRKHRLTCPCSDSYEKPKPEFLERFKSL 112
 DB 114 PPSCTPDYLMSTYVRGSDRNFKITPTIDLYNWTAPVQVFNCKAL 160

RESULT 7
 IDH_AZOVI STANDARD; PRT; 741 AA.
 AC P16100;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE ISOCITRATE DEHYDROGENASE [NADP] (EC 1.1.1.42) (OXALOSUCCINATE.
 DE DECARBOXYLASE) (IDH).
 GN ICD.
 OS Azotobacter vinelandii.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Azotobacter.
 OX NCBI_TaxID=354;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sahara T.; Suzuki M.; Tsuruba J.I.; Takada Y.; Abe K.; Fukunaga N.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP PRELIMINARY SEQUENCE OF 228-250 AND 254-259.
 RX MEDLINE=74086945; PubMed=4149369;
 RA Edwards D.J.; Heirlikson R.L.; Chung A.E.;
 RT "Triphosphopyridine nucleotide specific isocitrate dehydrogenase from
 RT Azotobacter vinelandii. Alkylation of a specific methionine residue
 RT and amino acid sequence of the peptide containing this residue.";
 RL Biochemistry 13:677-683(1974).
 CC -!- CATALYTIC ACTIVITY: ISOCITRATE + NADP(+) = 2-OXOGLUTARATE +
 CC CO(2) + NADPH.
 CC -!- ENZYME REGULATION: INHIBITION OF THIS ENZYME BY PHOSPHORYLATION
 CC REGULATES THE BRANCH POINT BETWEEN THE KREBS CYCLE AND THE
 CC GLYOXALATE BYPASS, WHICH IS AN ALTERNATE ROUTE THAT ACCUMULATES
 CC CARBON FOR BIOSYNTHESIS WHEN ACETATE IS THE SOLE CARBON SOURCE
 CC FOR GROWTH. THE PHOSPHORYLATION STATE OF THIS ENZYME IS CONTROLLED
 CC BY ISOCITRATE DEHYDROGENASE KINASE/PHOSPHATASE (ACEK).
 CC -!- SUBUNIT: MONOMER.
 CC -!- SIMILARITY: BELONGS TO THE MONOMERIC-TYPE FAMILY OF IDH.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D73443; BAA11169.1; -
 DR PIR: A10759; A10759.
 KW Oxidoreductase; NADP; Phosphorylation; Glyoxylate bypass;
 KW Tricarboxylic acid cycle.
 SQ SEQUENCE 741 AA; 80389 MW; 29FF352785AED8B CRC64;

Query Match 12.9%; Score 78; DB 1; Length 741;
 Best Local Similarity 24.8%; Pred. No. 3;
 Matches 26; Conservative 23; Mismatches 34; Indels 22; Gaps 4;

QY 2 DRHMRQLDIDVDLKNVNDL-----VPEFLPAPDEVETNCWSAFSCFOKAQLKS 55
 DB 77 DANIILKPNIASVPQKAAIKELQQQGYKLDPY---PEPKTDTDKDVRARYDKINGSA 133
 QY 56 AN-----TGNNERINVSIIKKLRKPPSTNAGRRQKHRLTCPSADS 96
 DB 134 VNPVREGNSDRAPLSVKVYARKHP-----HKMGWSADS 169

RESULT 8

ID ALB2_XENLA STANDARD; PRT; 607 AA.
 AC P14872;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE 74 KDA SERUM ALBUMIN PRECURSOR.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 [1]
 RN RP SEQUENCE OF 3-607 FROM N.A.
 RX MEDLINE=89313788; PubMed=2747653;
 RA Moskowitz J.E., Sargent T.D., Smith L.H. Jr., Pastori R.L.,
 RA Schoenberg D.R.;
 RT "Xenopus laevis serum albumin: sequence of the complementary
 RT deoxyribonucleic acids encoding the 68- and 74-kilodalton peptides
 RT and the regulation of albumin gene expression by thyroid hormone
 RT during development.";
 RL Mol. Endocrinol. 3:464-473(1989).
 [2]
 RN RP SEQUENCE OF 1-48 FROM N.A.
 RX MEDLINE=88172470; PubMed=2451026;
 RA Schorpp M., Doebebling U., Wagner U., Ryffel G.U.;
 RT "5'-flanking and 5'-proximal exon regions of the two xenopus albumin
 RT genes. Deletion analysis of constitutive promoter function.";
 RL J. Mol. Biol. 199:83-93(1988).
 [3]
 RN RP SEQUENCE OF 459-557 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=85126974; PubMed=3971963;
 RA Wolffe A.P., Glover J.F., Martin S.C., Tenniswood M.P.R.,
 RA Williams J.L., Tata J.R.;
 RT "Deinduction of transcription of Xenopus 74-kda albumin genes and
 RT destabilization of mRNA by estrogen in vivo and in hepatocyte
 RT cultures.";
 RL Eur. J. Biochem. 146:489-496(1985).

CC -!- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD
 CC BINDING CAPACITY FOR WATER, CA⁺⁺, NA⁺, K⁺, FATTY ACIDS, HORMONES,
 CC BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
 CC COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -!- TISSUE SPECIFICITY: PLASMA.
 CC -!- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.

CC -!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M21442; AAA49637.1; -
 DR EMBL: M28276; AAA49642.1; -
 DR PIR: B41882; ABXL72.
 DR HSP; P02768; IUOR.
 DR InterPro; IPR000264; -
 DR Pfam; PF00273; Transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR PROSITE; PS00212; ALBUMIN; 3.
 KW Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; Signal;
 KW Copper.

FT SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 24 POTENTIAL.
 FT CHAIN 25 607 74 KDA SERUM ALBUMIN.
 FT REPEAT 29 211 1.
 FT REPEAT 217 403 2.
 FT REPEAT 409 601 3.
 FT METAL 30 30 COPPER (BY SIMILARITY).
 FT DISULFID 80 88 BY SIMILARITY.
 FT DISULFID 101 117 BY SIMILARITY.
 FT DISULFID 116 127 BY SIMILARITY.
 FT DISULFID 147 192 BY SIMILARITY.
 FT DISULFID 191 200 BY SIMILARITY.
 FT DISULFID 223 269 BY SIMILARITY.
 FT DISULFID 268 276 BY SIMILARITY.
 FT DISULFID 288 302 BY SIMILARITY.
 FT DISULFID 301 312 BY SIMILARITY.
 FT DISULFID 339 384 BY SIMILARITY.
 FT DISULFID 383 392 BY SIMILARITY.
 FT DISULFID 415 461 BY SIMILARITY.
 FT DISULFID 460 471 BY SIMILARITY.
 FT DISULFID 484 500 BY SIMILARITY.
 FT DISULFID 499 510 BY SIMILARITY.
 FT DISULFID 537 582 BY SIMILARITY.
 FT DISULFID 581 590 BY SIMILARITY.
 FT CONFLICT 503 503 S -> L (IN REF. 3).
 FT CONFLICT 531 531 H -> D (IN REF. 3).
 SQ SEQUENCE 607 AA; 70382 MW; 592BA4177A36B66B CRC64;

Query Match 12.8%; Score 77.5; DB 1; Length 607;
 Best Local Similarity 24.8%; Pred. No. 2.7;
 Matches 35; Conservative 17; Mismatches 48; Indels 41; Gaps 5;

QY 13 DIVOLKNVNDLVPEF-----LPAPDEVETNCWSAFSCF-----OKAQLKSAN 57
 DB 72 BINDFAKSCINDKTPECEKPVGTLEFDKLCADPAVGNVYENSKCAKQDPERAOCKFAH 131

QY 58 TGNNERII-----NVSIIKKLRK-----PPSTNAGRRQKHRLTCPS 93

DB 132 RDHEHTSIRPEETCKLKEHPDILLSAFTHEARNHPDLYPPAVLTAUTKTHKLAHC 191

QY 94 CDSVEKKPKPEFLERFKSLQ 114

DB 192 CEEEDKE--KCFSEKMKQLMK 210

RESULT 9

ID IL15_PIG STANDARD; PRT; 162 AA.
 AC Q95253;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)

```
DE GN INTERLEUKIN-15 PRECURSOR (IL-15).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
ON NCBI_TaxID=9823;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=97449311; PubMed=9305780;
RA Canals A., Grimm D.R., Gasbarre L.C., Lunney J.K., Zarling D.S.;
RT "Molecular cloning of cDNA encoding porcine interleukin-15.";
RL Gene 195:337-339(1997).
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U58142; AAB72031.1; -
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48 POTENTIAL.
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 119 119 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 162 AA; 18437 MW; 7EF7992391893446 CRC64;

Query Match 12.6%; Score 76.5; DB 1; Length 162;
Best Local Similarity 24.1%; Pred. No. 0.8;
Matches 28; Conservative 19; Mismatches 42; Indels 27; Gaps 6;

Qy 14 IVDQLKNVNDLVPPE-----LPAPEDVETNCWAFSCF-----QKALKSANTG 59
Db 53 VISDLKK-IEDLIRSIHMDATLYTESDAHPNCKVTAMKCFLELRVILQESRNSDISDTV 111
Qy 60 NNERII-NVSTKKLRKPPSNAGRRQRHLTCPCSDSYEKKPKPEFLERFKSLIQ 114
Db 112 ENLIILANSSUSSIEYK---TESG-----CKECELEERKINEFLKSFTHIVQ 156

RESULT 10
ID IF2A_YEAST STANDARD; PRT; 304 AA.
AC P20459;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE EUKARYOTIC TRANSLATION INITIATION FACTOR 2 ALPHA SUBUNIT (EIF-2-
DE ALPHA).
GN SUF2 OR TIF211 OR YJR007W OR J1429.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89202411; PubMed=2649894;
RA Cigan A.M., Pabich E.K., Peng L., Donahue T.F.;
RT "Yeast translation initiation suppressor su12 encodes the alpha
RT subunit of eukaryotic initiation factor 2 and shares sequence
RT identity with the human alpha subunit.";
```

```
RL IL15.
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RA de Haan M., Smits P.H.M., Grivell L.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: EIF-2 FUNCTIONS IN THE EARLY STEPS OF PROTEIN SYNTHESIS
CC BY FORMING A TERNARY COMPLEX WITH GTP AND INITIATOR TRNA. THIS
CC COMPLEX BINDS TO A 40S RIBOSOMAL SUBUNIT, FOLLOWED BY MRNA BINDING
CC TO FORM A 43S PREINITIATION COMPLEX. JUNCTION OF THE 60S RIBOSOMAL
CC SUBUNIT TO FORM THE 80S INITIATION COMPLEX IS PRECEDED BY
CC HYDROLYSIS OF THE GTP BOUND TO EIF-2 AND RELEASE OF AN EIF-2-GDP
CC BINARY COMPLEX. IN ORDER FOR EIF-2 TO RECYCLE AND CATALYZE ANOTHER
CC ROUND OF INITIATION, THE GDP BOUND TO EIF-2 MUST EXCHANGE WITH GTP
CC BY WAY OF A REACTION CATALYZED BY EIF-2B.
CC -!- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA, AND A GAMMA
CC CHAIN.
CC -!- SIMILARITY: BELONGS TO THE EIF-2-ALPHA FAMILY.
CC -!- SIMILARITY: CONTAINS 1 'SI MOTIF' DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M25552; AAA70332.1; -
DR EMBL; X87611; CAA60929.1; -
DR EMBL; Z49507; CAA89529.1; -
DR PIR; A32108; A32108.
DR SGD; S0003767; SUI2.
DR InterPro; IPR003029; -
DR Pfam; PF00575; S1; 1.
KW Initiation factor; Protein biosynthesis; RNA-binding;
KW Phosphorylation.
FT MOD_RES 52 52 PHOSPHORYLATION (BY GCN2).
FT SEQUENCE 304 AA; 34717 MW; AF4F1C80303A4E98 CRC64;

Query Match 11.9%; Score 72; DB 1; Length 304;
Best Local Similarity 33.3%; Pred. No. 4.4;
Matches 20; Conservative 14; Mismatches 22; Indels 4; Gaps 4;

Qy 13 DIVDOLKNVND-LVPEFLPAPEDVETNC-EMSAFSCFQKALKSANTGNERNIIVSIK 70
Db 164 DVLDLKNVTSIKRLTPQAVKIRADVEVSCFSYEGIDAIKDA-LKSAEDMSTEQ-MQVKVK 221

RESULT 11
ID ALLA_YEAST STANDARD; PRT; 195 AA.
AC P32459;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE UREIDOGLYCOLATE HYDROLASE (EC 3.5.3.19).
GN DAL3 OR YIR032C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92133160; PubMed=1776360;
RA Yoo H.S., Cooper T.G.;
RT "The ureidoglycollate hydrolase (DAL3) gene in Saccharomycetes
RL cerevisiae."
RL Yeast 7:693-698(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
```

RA Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
 RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
 RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
 RA Walsh S.V., Whitehead S.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: UTILIZATION OF PURINES AS SECONDARY NITROGEN SOURCES,
 CC WHEN PRIMARY SOURCES ARE LIMITING.
 CC -!- CATALYTIC ACTIVITY: (-)-UREIDOGLYCOLATE + H(2)O = GLYOXYLATE +
 CC 2 NH(3) + CO(2).
 CC -!- PATHWAY: THIRD STEP IN THE DEGRADATION OF ALLANTOIN (PURINE
 CC CATABOLISM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M64778; AAA73025.1; -;
 DR EMBL: 238061; CA86192.1; -;
 DR PIR: S42022; S42022.
 DR PIR: S48494; S48494.
 DR SGD: S0001471; DAL3.
 KW Hydrolase; Purine metabolism; Prenylation; Lipoprotein.
 FT LIPID 192 192 FARNESYL (BY SIMILARITY).
 SQ SEQUENCE 195 AA; 21727 MW; 6230AEE59585206B CRC64;

 Query Match 11.8%; Score 71.5; DB 1; Length 195;
 Best Local Similarity 29.2%; Pred. No. 3;
 Matches 19; Conservative 9; Mismatches 24; Indels 13; Gaps 1;

 QY 15 VQOLKNVNDLVEFLPAPEDVETNCWAFSCFQAKLKSANTGNNRINVSIRKLKR 74
 Db 48 VSOVENKSTKVP-----NNWLFRCFQPHLNRFVETGNSQAISHSIRKLEK 94
 QY 75 KPPST 79
 Db 95 HPCST 99

 RESULT 12
 Y069 NPVAC
 ID Y069 NPVAC STANDARD; PRT; 262 AA.
 AC P41469;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE HYPOTHETICAL 30.4 KDA PROTEIN IN LEF3-IAP2 INTERGENIC REGION.
 OS Autographa californica nuclear polyhedrosis virus (ACMNPV).
 CC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 CC Nucleopolyhedrovirus.
 CC NCBI_TaxID=46015;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C6;
 RX MEDLINE=94303173; PubMed=8030224;
 RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
 RT "The complete DNA sequence of Autographa californica nuclear
 RL polyhedrosis virus.";
 RL Virology 202:586-605(1994).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL: L22858; AAA66699.1; -;
 DR InterPro: IPR002877; -;
 DR Pfam: PF01728; FtsJ; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 262 AA; 30355 MW; 05C1B44B00406AEC CRC64;

 Query Match 11.6%; Score 70; DB 1; Length 262;
 Best Local Similarity 22.2%; Pred. No. 5.9;
 Matches 18; Conservative 19; Mismatches 34; Indels 10; Gaps 2;

 QY 39 NCWSAFSCFQAKLKSANTGNNRINVSIRKLKRKPPSTNAGRRQKH-----RLTQPS 93
 Db 176 NCVLKVFDAFEHETIOMLN-----KFNHFKEKWLVIKPPSSRPANSEYRLICFNKLVRPY 230
 QY 94 CDSYEKKPKPEFLERFKSLIQ 114
 Db 231 CNYVNELEKOFERYRIQLK 251

 RESULT 13
 IFT3 MOUSE
 ID IFT3_MOUSE STANDARD; PRT; 403 AA.
 AC Q64345;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE INTERFERON-INDUCED PROTEIN WITH TETRATRICOPEPTIDE REPEATS 3 (IFT3-3)
 DE (GLUCOCORTICOID-ATTENUATED RESPONSE GENE 49 PROTEIN) (GARG-49) (IRG2).
 GS IFT3 OR IFT49 OR GARG49.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96239145; PubMed=8660659;
 RA Smith J.B., Herschman H.R.;
 RT "The glucocorticoid attenuated response genes GARG-16, GARG-39, and
 RT GARG-49/IRG2 encode inducible proteins containing multiple
 RT tetratricopeptide repeat domains.";
 RL Arch. Biochem. Biophys. 330:290-300(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C;
 RX MEDLINE=94267181; PubMed=8207206;
 RA Lee C.G.G., Demarqoy J., Jackson M.J., O'Brien W.E.;
 RT "Molecular cloning and characterization of a murine LPS-inducible
 RT cDNA.";
 RL J. Immunol. 152:5758-5767(1994).
 CC -!- SIMILARITY: BELONGS TO THE IFT FAMILY.
 CC -!- SIMILARITY: CONTAINS TPR REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U43086; AAC52627.1; -;
 DR EMBL: L32974; AAA39329.1; -;
 DR MGD; MGI:1101055; Ift3.
 KW Repeat; TPR repeat; Interferon induction.
 FT REPEAT 51 85 TPR 1.
 FT REPEAT 94 128 TPR 2.
 FT REPEAT 155 170 TPR 3B.
 FT REPEAT 173 207 TPR 4.
 FT REPEAT 208 241 TPR 5.
 FT REPEAT 242 275 TPR 6.
 FT REPEAT 276 286 TPR 7A.

[illegible]

Result No.	Score	Query		DB	Length	ID	Description
		Match	%				
1	95.5	15.8	114	4	Q0UBA3		Q0UBA3 homo sapien
2	95.5	15.8	136	4	Q00440		Q00440 homo sapien
3	85.5	14.1	162	6	Q9XSJ6		Q9XSJ6 ovnis aries
4	82	13.6	567	11	Q0S208		Q0S208 mus musculu
5	77.5	12.8	3052	14	Q82933		Q82933 johnsongras
6	75	12.4	566	5	Q9VYE8		Q9VYE8 drosophila
7	74	12.2	2197	5	Q9G296		Q9G296 plasmodium
8	73.5	12.1	868	5	Q9NAH8		Q9NAH8 caenorhabdi
9	72	11.9	187	13	Q9W756		Q9W756 gallus gall
10	72	11.9	257	1	Q27054		Q27054 methanobact
11	71	11.7	719	5	Q27365		Q27365 caenorhabdi
12	71	11.7	722	5	Q21448		Q21448 caenorhabdi
13	71	11.7	1101	5	Q62471		Q62471 caenorhabdi
14	70.5	11.7	848	4	Q75141		Q75141 homo sapien
15	70	11.6	262	14	Q92434		Q92434 bombyx mori
16	69	11.4	770	5	Q44014		Q44014 leishmania
17	69	11.4	1005	4	Q75336		Q75336 homo sapien
18	69	11.4	1220	10	Q24015		Q24015 lycopersico
19	68.5	11.3	336	11	Q62612		Q62612 rattus norv

```

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)
DE INTERLEUKIN-15.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Meazza R., Ferrini S.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y09908; CAA71044.1; -.
SQ SEQUENCE 136 AA; 15259 MW; 494ED8DDC5D364CD CRC64;

Query Match 15.8%; Score 95.5; DB 4; Length 136;
Best Local Similarity 25.9%; Pred. No. 0.014;
Matches 29; Conservative 22; Mismatches 46; Indels 15; Gaps 5;

QY 12 IDIVDQKYNVDLPPEF-----LPAPEDVETNCWAFSCF----KQAQLKSANTGNNE 62
Db 25 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVKTAMKCFLLLEQLVISLSEGDASIH 83
QY 63 RIINVSIIKKLKRKPPSTNAGRRKHRLTCPCSDSYEKPKPEFLERFKSLIQ 114
Db 84 TVENLIII--LANNSSNGNVTES---GCKEELEEKNIKEFLQSFVHIQV 130

RESULT 3
Q9XSJ6 ID Q9XSJ6 PRELIMINARY; PRT; 162 AA.
AC Q9XSJ6;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE INTERLEUKIN-15.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Casey G.J., Chaplin P.J.;
RT "Isolation of interleukin-15 mRNA transcripts from T and B cells
circulating in effluent lymph.";
RL J. Interferon Cytokine Res. 0:0-0(1999).
DR EMBL: AF149700; AAD37425.1; -.
SQ SEQUENCE 162 AA; 18510 MW; 87195650E45FLESD CRC64;

Query Match 14.1%; Score 85.5; DB 6; Length 162;
Best Local Similarity 28.8%; Pred. No. 0.18;
Matches 23; Conservative 10; Mismatches 46; Indels 1; Gaps 1;

QY 35 DVETNCWAFSCFQKQAQLKSANTGNNERIINVSIIKKLRKPPSTNAGRRKHRLTCPCSC 94
Db 78 DAHPNCKVTALQCFLLERLVILHESKNAAYVEI-IENLTMLADRLNLSSENKTELGCCKE 136
QY 95 DSYEKPKPEFLERFKSLIQ 114
Db 137 EELEKKSIIKEFLKSFVHIQV 156

RESULT 4
Q05208 ID Q05208 PRELIMINARY; PRT; 567 AA.
AC Q05208;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE ST2L PROTEIN PRECURSOR.

```

```

GN LY84.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93170492; PubMed=7916701;
RA Yanagisawa K., Takagi T., Tsukamoto T., Tetsuka T., Tominaga S.;
RT "Presence of a novel primary response gene ST2L, encoding a product
highly similar to the interleukin 1 receptor type 1.";
RL FEBS Lett. 318:83-87(1993).
DR EMBL: D13695; BA02854.1; -.
DR MGI: 98427; Ly84.
DR INTERPRO: IPR000157; -.
DR INTERPRO: IPR002052; -.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; Ig 3.
DR PFAM: PF01582; TIR; 1.
DR PROSITE: PS00092; N6_MTASE; UNKNOWN_1.
KW SIGNAL.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 567 ST2L PROTEIN.
SQ SEQUENCE 567 AA; 64801 MW; 9228CE27D95B0BC CRC64;

Query Match 13.6%; Score 82; DB 11; Length 567;
Best Local Similarity 27.1%; Pred. No. 1.5;
Matches 29; Conservative 17; Mismatches 29; Indels 32; Gaps 6;

QY 16 DQLKNVNDLPPEFLPAPEDVETNCWAFSCFQKQAQLKSANTGNNERIINVSIIKKLRK 75
Db 76 DRLK-----FLPARVE-----DSGIYACV----IRSPNL-NRTGYLNVTHK---K 113
QY 76 PPSTN-----AGRRQKHRLTCPCSDSYEKPKPEFLERFKSL 112
Db 114 PPSCNIPDYLMTYVTRGSDKNFKITCTIDLYNWTAPVQWFKCKAL 160

RESULT 5
Q82933 ID Q82933 PRELIMINARY; PRT; 3052 AA.
AC Q82933;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE JOHNSON GRASS MOSAIC VIRUS PROTEASE 1 AND 3, HELPER COMPONENT 6K
PROTEIN, COAT PROTEIN, NUCLEAR INCLUSION PROTEINS.
OS Johnsongrass mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=31742;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94200956; PubMed=8150599;
RA Gough K.H., Shukla D.D.;
RT "Nucleotide sequence of Johnsongrass mosaic potyvirus genomic RNA.";
RL Intervirology 36:181-192(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Gough K.H.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Nurchayati E.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z26920; CAA81549.1; -.
DR MEROPS; C04.001; -.
DR MEROPS; C06.001; -.
DR MEROPS; S30.001; -.
DR INTERPRO: IPR001205; -.
DR INTERPRO: IPR001410; -.
DR INTERPRO: IPR001456; -.

```

DR INTERPRO: IPR001592; -
 DR INTERPRO: IPR001730; -
 DR INTERPRO: IPR002540; -
 DR PFAM: PF00680; RNA_dep_RNA_pol; 1.
 DR PFAM: PF00767; Poly_coat; 1.
 DR PFAM: PF00851; Peptidase_C6; 1.
 DR PFAM: PF00863; Peptidase_C4; 1.
 DR PFAM: PF01577; Poly_P1; 1.
 DR PRINTS: PR00966; NIAPOTYPASE.
 KW Coat protein; Protease.
 FT CHAIN 239 699 HELPER COMPONENT (HC).
 FT CHAIN 700 1046 PROTEASE3.
 FT CHAIN 1047 1098 6K1.
 FT CHAIN 1099 1757 CYTOPLASMIC INCLUSION (CI).
 FT CHAIN 1758 1803 6K2.
 FT CHAIN 1804 1991 SMALL NUCLEAR INCLUSION/VIRUS ATTACHED
 FT CHAIN 1992 2233 PROTEIN.
 FT CHAIN 2234 2749 SMALL NUCLEAR INCLUSION/PROTEASE.
 FT CHAIN 2750 3052 LARGE NUCLEAR INCLUSION/POLYMERASE.
 FT CHAIN 3052 34729 MW; 08CD8831A73EBCA9 CRC64;
 SQ SEQUENCE 3052 AA; 34729 MW; 08CD8831A73EBCA9 CRC64;
 Query Match 12.8%; Score 77.5; DB 14; Length 3052;
 Best Local Similarity 23.5%; Pred. No. 27;
 Matches 24; Conservative 20; Mismatches 55; Indels 3; Gaps 3;
 QY 14 IVDLKYNVDLPPELPAPEDVETNCWSAFSCFQAOLKSANTGNRIINVSIIKLK 73
 DB 220 LVNALQYED-VKQICHVSFDEARAFWKGFTENTTAQRREAHDTNHEPV-MSVEECG 277
 QY 74 RKPPSTNAGRKRLTSC-DSYEKKPKPEFLERFKSLQ 114
 DB 278 RRAAMLENAPHQGFKTCRHCFTDEHSDSEVCERHNAIQ 319
 RESULT 6
 QYVYEB PRELIMINARY; PRT; 566 AA.
 AC QYVYEB8
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE CG15745 PROTEIN.
 GN CG15745.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; "
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003492; AAF48250.1; -
 DR FLYBASE; FBgn0030469; CG15745.
 SQ SEQUENCE 566 AA; 60620 MW; 31FAA1FA4AEF427F CRC64;
 Query Match 12.4%; Score 75; DB 5; Length 566;
 Best Local Similarity 24.7%; Pred. No. 8.3;
 Matches 20; Conservative 19; Mismatches 36; Indels 6; Gaps 2;
 QY 35 DVETNCWSAFS-CFQAOLKSANTGNRIINVSIIKLKRPSTNAGRKRLTCTP 92
 DB 120 DIGPCEVASISELAKRAQLKAQFFGNQVG---GLARDSETSTTRITNTNRSAYP 175
 QY 93 SCDSEYKPKPKPEFLERFKSL 113
 DB 176 SKTERGKPVQQLIDQFQMI 196
 RESULT 7
 ID 096296 PRELIMINARY; PRT; 2197 AA.
 AC 096296
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
 DE PTEMPL
 GN PEB1055C.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99021743; PubMed=9804551;
 RA Gardner M.J., Tettelin H., Carucci D.J., Cummings L.M., Aravind L.,
 RA Koonin E.V., Shallom S., Mason T., Yu K., Fujii C., Pederson J.,
 RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M.,
 RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
 RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
 RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
 falciparum."
 RL Science 282:1126-1132(1998).
 DR EMBL; AE001434; AAC71996.1; -
 SQ SEQUENCE 2197 AA; 249668 MW; D2A92DB54535C143 CRC64;
 Query Match 12.2%; Score 74; DB 5; Length 2197;
 Best Local Similarity 29.5%; Pred. No. 44;
 Matches 28; Conservative 16; Mismatches 41; Indels 10; Gaps 4;
 QY 18 LKNVNDLVPELPAPEDVETNCWSAF--SCFQAOLKSANTGNRIINVSIIKLK 75
 DB 1594 VKSFLETWIPK-IAVVDQDNVILSKFGNSGCSASAISTN-GNEIDAIDCMIKRLK 1651

QY 76 -----PPSTNAGRRQKRLCTPCSDSYEKPKPK 104
Db 1652 IDECKRKPGNSGQTNETLTHPLDVQDEPLEE 1686

RESULT 8

Q9NAH8 PRELIMINARY; PRT; 868 AA.
AC Q9NAH8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE Y47D3A.14 PROTEIN.
GN Y47D3A.14.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Matthews L.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology".
RL Science 282:2012-2018(1998).
DR EMBL; AL117202; CAB55073.1; -
SQ SEQUENCE 868 AA; 100165 MW; 09A2A97066E373D7 CRC64;

Query Match 12.1%; Score 73.5; DB 5; Length 868;
Best Local Similarity 29.2%; Pred. No. 19;
Matches 19; Conservative 11; Mismatches 30; Indels 5; Gaps 1;

QY 42 WSAFSCFOAKLKS-----ANTGNNERIINVSIIKKLRKPPSTNAGRRQKRLCTPCSDS 96
Db 680 WSRVQAFKKNKMSGGGTNSDSSEKSEKVLKLRKPPSTNAGRRQKRLCTPCSDS 96

QY 97 YEKPK 101

Db 740 HVAIP 744

RESULT 9

Q9W756 PRELIMINARY; PRT; 187 AA.
AC Q9W756;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (INTERLEUKIN 15 PRECURSOR).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCWL; TISSUE=LIVER;
RA Burnside J.; Sofer L.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SC; TISSUE=LIVER;
RA Choi K.D.; Lillehoj H.S.; Burnside J.;
RL "Gallus gallus mRNA for IL-15 precursor".
DR EMBL; AF152927; AAD38392.1; -
DR EMBL; AF139097; AAF61446.1; -
SQ SEQUENCE 187 AA; 21964 MW; 470601EBF8837095 CRC64;

Query Match 11.9%; Score 72; DB 13; Length 187;
Best Local Similarity 22.7%; Pred. No. 5.3;
Matches 29; Conservative 15; Mismatches 46; Indels 38; Gaps 5;

QY 1 QDRMIRMROLIDIVDOLKNVNDLVPEFLPAPEDVETNCWSAFSCF---QKAOLKSN 57
Db 76 KDLLEIKTSEDIDVSLYAN-----TYEDIE--COEPMRCFFLEMKVILHECD 122
QY 58 TGNNERIINVSIIKKLRKPPSTNAGRRQKRLCTPCSDSYEKPKPKPK 106
Db 123 -----IKKSRKHVDVRNLIWKNAGNFATYQLNSTTAKKCKECEEENFTFEI 171
QY 107 EREKSLLO 114
Db 172 QSEVKVIQ 179

RESULT 10

O27054 PRELIMINARY; PRT; 257 AA.
AC O27054;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE CONSERVED PROTEIN.
GN MTH973.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanobacterium.
OX NCBI_TaxID=2166;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R.; Doucette-Stamm L.A.; DeLoughery C.; Lee H.-M.; DuBois J.;
RA Aldredge T.; Bashirzadeh R.; Blakely D.; Cook R.; Gilbert R.;
RA Harrison D.; Hoang L.; Keagie P.; Lumm W.; Pothier B.; Qiu D.;
RA Spadafora R.; Vicare R.; Wang Y.; Wierzbowski J.; Gibson R.;
RA Jiwni N.; Caruso A.; Bush D.; Safer H.; Patwell D.; Prabhakar S.;
RA McDougall S.; Shimer G.; Goyal A.; Pietrovski S.; Church G.M.;
RA Daniels C.J.; Mao J.-I.; Rice P.; Nolling J.; Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics".
RL J. Bacteriol. 179:7135-7155(1997).
DR EMBL; AE000871; AAB85469.1; -
DR INTERPRO; IPR002811; -
DR PFAM; PF01958; DUF108; 1.
DR PRODOM; PD017325; -; 1.
SQ SEQUENCE 257 AA; 27569 MW; D9A387A535611FC7 CRC64;

Query Match 11.9%; Score 72; DB 1; Length 257;
Best Local Similarity 25.0%; Pred. No. 7.4;
Matches 27; Conservative 12; Mismatches 29; Indels 40; Gaps 3;

QY 5 MIRMROLIDIVDOLKNY-----VNDLVPEFLPAPEDV----- 36
Db 48 MVDGRAVLVDVADMLPEVDLVVEAASPEAVRDVPEILEAGKDVVMSVGLMDPELREML 107
QY 37 -----ETNCWSAFSCFOAKLKSANTGNNERIINVSIIKKLRKPP 77
Db 108 VELASLNDATIHVPSGAIVGLDGLKAASMGNI-----SVKLITRKPP 150

RESULT 11

Q27365 PRELIMINARY; PRT; 719 AA.
AC Q27365;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JAN-1999 (Tremblrel. 09, Last annotation update)
DE LIN-15 PROTEIN.

LN-15.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2; PubMed=7982579;
 RX MEDLINE=95073612; PubMed=7982579;
 RA Clark S.G., Lu X., Horvitz H.R.;
 RT "The Caenorhabditis elegans locus lin-15, a negative regulator of a
 tyrosine kinase signaling pathway, encodes two different proteins.";
 RL Genetics 137:987-997(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kershaw J.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U10413; AAA20090.1; -;
 DR EMBL; U10411; AAA20087.1; -;
 DR EMBL; Z79605; CAB01904.1; -;
 SQ SEQUENCE 719 AA; 80062 MW; 8E69CF00D419DD7F CRC64;

Query Match 11.7%; Score 71; DB 5; Length 719;
 Best Local Similarity 23.0%; Pred. No. 28;
 Matches 31; Conservative 17; Mismatches 43; Indels 44; Gaps 5;
 QY 3 RHIMRMQLDIVD-----QLKNVNDLVPEFLPAPED----- 35
 DB 255 RHE-RLRMCYDHDVDFVYEMCMDFAEATESEINEMPPDRLMRGHDIYALKRVGDLHKGK 313
 QY 36 VETNCWSAFSCFQAKLKSANTGNNRIINVSIIK-----LKRKPPSTNAGR 83
 DB 314 VTSNTPLYSF-----KNSIKSYRHNVPWVNGSLSKPKPSFELVALLOQSPSTNLINE 369
 QY 84 RQKHLCTPCSDSYE 98
 DB 370 LLNHNLSLSDADKQE 384

RESULT 12
 Q21448
 ID Q21448 PRELIMINARY; PRT; 722 AA.
 AC Q21448;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE LIN-15A.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-N2;
 RX MEDLINE=94331780; PubMed=8054684;
 RA Huang L.S., Tzou P., Sternberg P.W.;
 RT "The lin-15 locus encodes two negative regulators of Caenorhabditis
 elegans vulval development.";
 RL Mol. Biol. Cell 5:395-411(1994).
 DR EMBL; Z29967; CAA82855.1; -;
 SQ SEQUENCE 722 AA; 80893 MW; 7CDE136694DFA289 CRC64;

Query Match 11.7%; Score 71; DB 5; Length 722;
 Best Local Similarity 23.0%; Pred. No. 28;
 Matches 31; Conservative 17; Mismatches 43; Indels 44; Gaps 5;
 QY 3 RHIMRMQLDIVD-----QLKNVNDLVPEFLPAPED----- 35
 DB 255 RHE-RLRMCYDHDVDFVYEMCMDFAEATESEINEMPPDRLMRGHDIYALKRVGDLHKGK 313
 QY 36 VETNCWSAFSCFQAKLKSANTGNNRIINVSIIK-----LKRKPPSTNAGR 83

Db 314 VTSNTPLYSF-----KNSIKSYRHNVPWVNGSLSKPKPSFELVALLOQSPSTNLINE 369
 QY 84 RQKHLCTPCSDSYE 98
 DB 370 LLNHNLSLSDADKQE 384
 RESULT 13
 O62471
 ID O62471 PRELIMINARY; PRT; 1101 AA.
 AC O62471;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE Y45F10B.10 PROTEIN.
 GN Y45F10B.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McMurray A.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
 RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL Nature 368:32-38(1994).
 DR EMBL; AL021487; CAA16357.1; -;
 DR INTERPRO; IPR001680; -;
 DR PFAM; PF00400; WD40; 9.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR PROSITE; PS00678; WD_REPEATS; UNKNOWN 2.
 SQ SEQUENCE 1101 AA; 121014 MW; 9C9C657F8A67D6B0 CRC64;

Query Match 11.7%; Score 71; DB 5; Length 1101;
 Best Local Similarity 31.8%; Pred. No. 43;
 Matches 21; Conservative 12; Mismatches 23; Indels 10; Gaps 3;
 QY 49 QKAQLKSANTGNNRIINVSIIKLRKP--PSTNAGRROK-HRLTGPCSDSYEKKPKPEF 105
 DB 1013 RSARQSVSSASNEPVAESAGEIKKDKPTLSSNNGNAQSPRATAP-----KPTFDM 1065
 QY 106 LERKFS 111
 DB 1066 LERSKS 1071

RESULT 14
 O75141
 ID O75141 PRELIMINARY; PRT; 848 AA.
 AC O75141;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE KIAA0650 PROTEIN (FRAGMENT).
 GN KIAA0650.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Search completed: May 23, 2001, 11:19:43
Job time: 596 sec

```

Query Match      11.7% ; Score 70.5 ; DB 4 ; Length 848 ;
Best Local Similarity 25.9% ; Pred. No. 37 ;
Matches 21 ; Conservative 17 ; Mismatches 38 ; Indels 5 ; Gaps 3 ;

QY      28  EFIPAPEDVETNC-ENSAFSCFOKAQKLSANTGNNERIINVSIKKLRKPPSTNAGRQK  86
      . : : | | : : | | | | | | : : : : : : | | : : | | : :
Db      66  KFIPGGPGNKDLCTWFESDFIRVOLIS---GPPAKLLIDWPFLKESIPVIN-GRDLQ  121

QY      87  HRLTCPSCDSEYKKPPKEFLE  107

Db     122  NPITIVOLCDQNDNPAPVQHVK  142

```

```

Query Match      11.6%; Score 70; DB 14; Length 262;
Best Local Similarity 22.2%; Pred. No. 12;
Matches 18; Conservative 18; Mismatches 35; Indels 10; Gaps 2;

OY   39 NCEWAFSCFOAKLQSANTGNERNRIINVSIKKILKRPPSTNAGRRORKH-----RLTCPS 93
    || | : : : | : : | : | : | : | : | : | : | : | : | : | : |
DB   176 NCVLKVFDAFEHKTQMLN-----KFYNHFKEKWLYVPPSSRPANSERYLICFNKLVRPY 230

OY   94 CDSYEKKPKPEFLERFKSLQ 114
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:11:34 ; Search time 108.07 Seconds
(without alignments)
64.531 Million cell updates/sec

Title: US-09-522-217-2_COPY_41_162
Perfect score: 647
Sequence: 1 QLIDIVDLKNVNDLPEF.....LLQKMIHQHLSRTHGSEDS 122

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues 390729
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	647	100.0	162	21	B18623 A human zalphall 1
2	647	100.0	519	21	B18627 Amino acid sequenc
3	368	56.9	146	21	B18624 A mouse zalphall 1
4	368	56.9	510	21	B18628 Amino acid sequenc
5	176	27.2	32	21	B18626 Antigeninc peptide
6	164	25.3	40	21	B18625 Antigeninc peptide
7	103.5	16.0	135	21	B18625 Human Interleukin-
8	99.5	15.4	114	16	R83310 Human Interleukin-
9	99.5	15.4	114	16	R83435 Mammalian interleu
10	99.5	15.4	114	16	R66928 Mammalian IL-15.
11	99.5	15.4	114	17	W09099 Simian mature epit

12	99.5	15.4	114	17	W07253 Generic mammalian
13	99.5	15.4	114	19	W39186 Simian epithelium
14	99.5	15.4	114	20	W39186 Simian epithelium
15	99.5	15.4	114	21	Y52309 Mature simian epit
16	99.5	15.4	162	16	R83438 Human interleukin-
17	99.5	15.4	162	16	R66927 Human IL-15. Homo
18	99.5	15.4	162	17	W09098 Simian epithelium
19	99.5	15.4	162	17	W07255 Human epithelium-d
20	99.5	15.4	162	17	R98527 Human interleukin-
21	99.5	15.4	162	18	W37369 Wild-type interleu
22	99.5	15.4	162	18	W01658 Human interleukin-
23	99.5	15.4	162	19	W53878 Human interleukin-
24	99.5	15.4	162	19	W39185 Simian epithelium
25	99.5	15.4	162	20	Y03756 Simian epithelium-
26	99.5	15.4	162	21	B18632 Amino acid sequenc
27	99.5	15.4	162	21	Y78595 Human interleukin-
28	99.5	15.4	162	21	Y52308 Simian epithelium-
29	99.5	15.4	162	22	B50870 Human IL-15. Homo
30	94.5	14.6	114	16	R83309 Simian interleukin
31	94.5	14.6	114	17	W09101 Human mature epit
32	94.5	14.6	114	19	W39188 Human epithelium d
33	94.5	14.6	114	20	Y03759 Human epithelium-d
34	94.5	14.6	114	21	Y52311 Mature human epit
35	94.5	14.6	122	17	R90842 Recombinant flag s
36	94.5	14.6	162	16	R83436 Simian interleukin
37	94.5	14.6	162	16	R66926 Simian IL-15. Cer
38	94.5	14.6	162	17	W09100 Human epithelium d
39	94.5	14.6	162	17	W07254 Simian epithelium-
40	94.5	14.6	162	17	R98526 Simian interleukin
41	94.5	14.6	162	17	R92798 Mammalian interleu
42	94.5	14.6	162	18	W37370 Mutant interleukin
43	94.5	14.6	162	19	W39187 Human epithelium d
44	94.5	14.6	162	20	Y03758 Human epithelium-d
45	94.5	14.6	162	21	Y78594 Simian interleukin

ALIGNMENTS

RESULT 1

B18623	ID	B18623 standard; Protein; 162 AA.
XX	AC	B18623;
XX	DT	22-JAN-2001 (first entry)
XX	DE	A human zalphall ligand polypeptide.
XX	DE	zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
KW	KW	tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
XX	OS	Homo sapiens.
XX	PN	WO200053761-A2.
XX	PD	14-SEP-2000.
XX	PF	09-MAR-2000; 2000MO-US060607.
XX	PR	09-MAR-1999; 99US-0264908.
PR	PR	11-MAR-1999; 99US-0265992.
PR	PR	01-JUL-1999; 99US-0142013.
XX	XX	(ZYMO) ZYMOGENETICS INC..
PA	PI	Novak JE, Presnell SR, Sprecher CA, Foster DC, Gross JA;
XX	PI	Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX	XX	WPI; 2000-565600/52.
DR	DR	N-PSDB; A75552.
XX	PT	New human cytokine, designated zalphall ligand, useful for stimulating

PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 PS Disclosure; Page 205-206; 256pp; English.
 XX The present sequence represents a human zalphall ligand polypeptide,
 CC which is a cytokine. The zalphall ligand is useful for stimulating the
 CC proliferation and development of haematopoietic cells in vitro and in
 CC vivo. Zalphall ligand polynucleotides can be used as primers or probes
 CC for cloning the zalphall gene. The zalphall ligand is useful for
 CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
 CC used for treating leukaemias and lymphomas. Antagonists against zalphall
 CC ligand are useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX Sequence 162 AA;
 SQ

Query Match 100.0%; Score 647; DB 21; Length 162;
 Best Local Similarity 100.0%; Pred. No. 2.3e-66;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QLIDIVDLKNVNDLVPEFLPAPEDVETNCWSAFSCFOAKLKSANTGNNERIINYSI 60
 DB 41 qlidivdlknyvndlvpeflpapedvetncwsafscfqaklsantgnneriinysi 100
 QY 61 KKLKRPSTNAGRQKRLTCPCSDSYEKKPKPEFLERFKSLQKMIHQHLSRTHGSE 120
 DB 101 kklrkppstnagrqrkrltccpscdsyekppkeflerfklsllqkmihqlsrrthgse 160
 QY 121 DS 122
 DB 161 ds 162

RESULT 2
 B18627
 ID B18627 standard; Protein; 519 AA.
 XX B18627;
 XX
 XX 22-JAN-2001 (first entry)
 DE Amino acid sequence of MBP-human zalphall ligand fusion in pTAP126.
 XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX Synthetic.
 OS Homo sapiens.
 XX WO2000053761-A2.
 XX 14-SEP-2000.
 XX 09-MAR-2000; 2000WO-US06067.
 XX 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX (ZYMO) ZYMOGENETICS INC.
 FA Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 XX

DR WPI: 2000-565600/52.
 DR N-PSDB; A75599.
 XX New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX Example 31; Page 233-235; 256pp; English.
 PS The present sequence represents a MBP-human zalphall ligand fusion in
 XX the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
 CC is useful for stimulating the proliferation and development of
 CC haematopoietic cells in vitro and in vivo. Zalphall ligand
 CC polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX Sequence 519 AA;
 SQ

Query Match 100.0%; Score 647; DB 21; Length 519;
 Best Local Similarity 100.0%; Pred. No. 1e-65;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QLIDIVDLKNVNDLVPEFLPAPEDVETNCWSAFSCFOAKLKSANTGNNERIINYSI 60
 DB 398 qlidivdlknyvndlvpeflpapedvetncwsafscfqaklsantgnneriinysi 457
 QY 61 KKLKRPSTNAGRQKRLTCPCSDSYEKKPKPEFLERFKSLQKMIHQHLSRTHGSE 120
 DB 458 kklrkppstnagrqrkrltccpscdsyekppkeflerfklsllqkmihqlsrrthgse 517
 QY 121 DS 122
 DB 518 ds 519

RESULT 3
 B18624
 ID B18624 standard; Protein; 146 AA.
 XX B18624;
 XX 22-JAN-2001 (first entry)
 DE A mouse zalphall ligand polypeptide.
 XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX Mus musculus.
 OS WO2000053761-A2.
 XX 14-SEP-2000.
 XX 09-MAR-2000; 2000WO-US06067.
 XX 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX (ZYMO) ZYMOGENETICS INC.
 PA

XX PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 XX WPI; 2000-565600/52.
 XX DR
 XX PT New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX Example 34; Page 227; 256pp; English.
 XX PS
 XX CC The present sequence was used to raise antibodies, and is derived from
 CC a human zalphall ligand polypeptide. Zalphall ligand is a cytokine.
 CC The zalphall ligand is useful for stimulating the proliferation and
 CC development of haematopoietic cells in vitro and in vivo. Zalphall
 CC ligand polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-aporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX Sequence 32 AA;
 SQ
 Query Match 27.2%; Score 176; DB 21; Length 32;
 Best Local Similarity 100.0%; Pred. No. 2.6e-13;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 82 CPSCDSYEKKPKPEFLERFKSLQKMIHQHLS 113
 DB 1 cpscdsyekppkeflerfklsllqkmibqhis 32
 RESULT 6
 B18625
 ID B18625 standard; Peptide; 40 AA.
 XX AC B18625;
 XX AC
 XX DT 22-JAN-2001 (first entry)
 XX DE Antigeninc peptide derived from a human zalphall ligand polypeptide.
 XX KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopolesis; B cell tumour.
 XX OS Homo sapiens.
 XX PN WO200053761-A2.
 XX PD 14-SEP-2000.
 XX PF 09-MAR-2000; 2000WO-US06067.
 XX PR 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX (ZYMO) ZYMOGENETICS INC.
 XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 XX WPI; 2000-565600/52.
 XX DR

XX PT New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX Example 34; Page 227; 256pp; English.
 XX PS
 XX CC The present sequence was used to raise antibodies, and is derived from
 CC a human zalphall ligand polypeptide. Zalphall ligand is a cytokine.
 CC The zalphall ligand is useful for stimulating the proliferation and
 CC development of haematopoietic cells in vitro and in vivo. Zalphall
 CC ligand polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-aporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX Sequence 40 AA;
 SQ
 Query Match 25.3%; Score 164; DB 21; Length 40;
 Best Local Similarity 100.0%; Pred. No. 8.1e-12;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QLIDIVDQLKYNVNDLVPEFLPAPEDVETNC 31
 DB 10 qlidivdqlknyvndlvpeflpapedvetnc 40
 RESULT 7
 Y54825
 ID Y54825 standard; Protein; 135 AA.
 XX AC Y54825;
 XX AC
 XX DT 04-FEB-2000 (first entry)
 XX DE Human Interleukin-15 protein sequence.
 XX KW Interleukin-15; IL-15; antisense inhibitor; untranslated region; 5'UTR;
 KW 3'UTR; human; infection; inflammation; tumour; therapy; diagnosis.
 XX OS Homo sapiens.
 XX PN US5985663-A.
 XX PD 16-NOV-1999.
 XX PF 25-NOV-1998; 98US-0200141.
 XX PR 25-NOV-1998; 98US-0200141.
 XX (ISIS-) ISIS PHARM INC.
 XX Bennett CF, Cowsert LM;
 PI WPI; 2000-022283/02.
 DR N-PSDB; 237358.
 XX Antisense compound useful for inhibiting human interleukin-15
 PT expression useful for treating diseases associated with interleukin-15
 PT expression -
 XX Example 13; Column 43-44; 31pp; English.
 XX DR

CC encoding mature simian or human IL-15 polypeptides. The active
CC simian and human IL-15 polypeptides are disclosed in R83309 &
CC R83310 respectively. The invention also comprises other mammalian
CC IL-15, including human IL-15, that hybridise to probes defined by
CC R83438. A plasmid contg. a recombinant clone of human IL-15
CC cDNA was deposited with ATCC on Feb. 19 1993 under ATCC 69245.
CC The deposit was named 141-hETf. R83435 is a mammalian mature
CC IL-15 polypeptide. It is a generic sequence which encompasses both
CC R83309 (simian) and R83310 (human) IL-15 mature polypeptides.
XX
SQ Sequence 114 AA:

```

Query Match      15.4%; Score 99.5; DB 16; Length 114;
Best Local Similarity 25.2%; Pred. No. 0.00075;
Matches 30; Conservative 23; Mismatches 45; Indels 21; Gaps 5;

QY 3 IDIVDQLKNYNDLVPEF-----LPAPEDVTNCEWSAFSCF-QKAQKLSANTGNN----- 52
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 nvnsldlkk-ledliqmhdatlytsdvhpscvktamkcflllelqvixesgdxihd 61

QY 53 --ERLIINVS TKKLRKPPSTNAGRQRHRLTQPCSDSYEKKPPPEFLRFRVSLSLQKMTIH 109
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 tvenliilannxlssnqnxtesq-----ckeceeeleekniflsgfvhiwcmfin 112

```

RESULT	10
R66928	
ID	R66928 standard; Protein; 114 AA.
XX	
XX	R66928;
XX	
DT	04-SEP-1995 (first entry)
XX	
XX	Mammalian IL-15.
DE	
XX	
KW	Interleukin-15; IL-15; T-cell growth factor;
KW	antitumor; virucide.
XX	
OS	Mammalia.

Key	Location/Qualifiers
FT	Misc-difference 52
FT	/label= Leu, His
FT	Misc-difference 57
FT	/label= Ala, Thr
FT	Misc-difference 58
FT	/label= Ser, Asp
FT	Misc-difference 73
FT	/label= Ser, Ile
FT	Misc-difference 80
FT	/label= Val, Ile

XX	2A9402636-A.	
PN		
XX		
XX	28-DEC-1994.	
PD		
XX		
XX	18-APR-1994; 94ZA-0002636.	
XX		
XX		
PR	18-APR-1994; 94ZA-0002636.	
XX		
XX	(IMMV) IMMUNEX CORP.	
PA		
XX		
XX	Anderson DM, Eisenman JR, Fung V, Grabstein KH;	
PI	Rauch C;	
PI		

XX WPI; 1995-082473/11.
DR
XX
XX New purified interleukin-15 - which induces T cell proliferation
PT and differentiation, used for the treatment of tumours and viral
PT infection
PT
XX
XX Claim 1; Page 33; 47pp; English.
PS

Claim 1; Page 33; 47pp; English.

```

XX Simian and human IL-15 cDNAs (Q84583-84) can be used to obtain
CC cDNAs encoding other mammalian homologs of IL-15. A general
CC sequence of mammalian IL-15 is claimed.
XX
SQ Sequence 114 AA;
      15.4%; Score 99.5; DB 16; Length 114;
Query Match Best Local Similarity 25.2%; Pred. No. 0.00075;
Matches 30; Conservative 23; Mismatches 45; Indels 21; Gaps 5;
Oy 3 IDIVDLKNNVNDLVLPPEF-----LPAPDEVETNCWSAFCF-QKAOLKSANTGNN---- 52
    : :: | | : | | : | : | : | : | : | : | : | : | : | : | :
Db 3 vnvvisdikk-tiedliqsmhidatlytesdvhpsscvtamkcflllelqvissesgdxihd 61
Oy 53 --ERIINVSIIKKLRKPSTNAGRQRKHRLTCPSCDSYEKKPKPKEFLERFKSLLOKMTH 109
    | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 62 tvenlililannxlssnguxxtesg-----ckeceeleeknikfqlsqfvhiwgmfin 112

```

RESULT 11
W09099
ID W09099 standard; Protein; 114 AA.

XX	W09099;
AC	
XX	
XX	11-MAR-1997 (first entry)
DT	
XX	
XX	Simlan mature epithelium derived T cell factor.
DE	
XX	
XX	setF; African green monkey; CV1/EBNA cell; T-cell; B-cell;
KW	lymphocyte; proliferation; differentiation; gastrointestinal;
KW	HIV infection; human immunodeficiency virus.
XX	
OS	Cercopithecus aethiops.

XX	Key	Location/Qualifiers
FH	Protein	1..114
FT		/label= mature_setf
FT		
PN	US5574138-A.	
PN		
XX		
XX	12-NOV-1996.	
PD		
XX		
XX	08-MAR-1993;	93US-0031399.
PF		
XX		
XX		
PR	22-FEB-1995;	95US-0393305.
PR	08-MAR-1995;	93US-0031399.
PR	22-APR-1994;	94US-0233606.
XX		
XX	(IMNV) IMMUNEX CORP.	
PA		

PI Anderson DM, Eisenman JR, Fung V, Grabstein KH;
PI Rauch C;

AA
DR WPI; 1996-517923/51.
DR N-PSDB: T49455.

New epithelium derived τ cell factor - induces proliferation of τ PT
PT and B cells, stimulates destruction of tumour and virus-infected
PT cells and protects against toxicity, partic. for treating intestinal
PT disease and HIV infection

PS Claim 1; Fig 1; 35pp; English.

CC The simian ETF (epithelium derived T cell factor) was isolated from
CC African green monkey CV1/EBNA cell conditioned medium. The N-
CC terminal sequence of the purified ETF was determined and then PCR
CC primers were designed based on the sequence information. A 92 bp
CC fragment was amplified from CV1/EBNA DNA and was used as a probe to
CC screen a CV1/EBNA cDNA library for the full-length ETF coding
CC sequence. Mature ETF induces proliferation and/or differentiation
CC

Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVTKAMKCFLELLEQLQVLSLESGDASIH 61
Qy 54 RIINVSIKKLRKPPSTNAGRQKRLTCPCSDSYEKPKPEFLERFKSLQKMIH 109
Db 62 TVENLII--LANNSSNGNVTES---GCKECEELEEKNIKEFLQSFVHVQMFN 112

RESULT 2

US-08-031-399-12
; Sequence 12, Application US/08031399
; Patent No. 5552303
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Epithelium-derived T-cell factor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/031.399
; FILING DATE: 19930308
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-031-399-12

Query Match 15.4%; Score 99.5; DB 1; Length 114;
Best Local Similarity 25.2%; Pred. No. 0.00025;
Matches 30; Conservative 23; Mismatches 45; Indels 21; Gaps 5;

Qy 3 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCESAFSCF-QKAOLKSANTGN--- 52
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVTKAMKCFLELLEQLQVLSLESGDASIH 61
Qy 53 --ERIINVSIKKLRKPPSTNAGRQKRLTCPCSDSYEKPKPEFLERFKSLQKMIH 109
Db 62 TVENLIIANNXNLSNGNXTESG-----CKECEELEEKNIKEFLQSFVHVQMFN 112

RESULT 3

US-08-393-305-3
; Sequence 3, Application US/08393305
; Patent No. 5574138
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor

; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393.305
; FILING DATE: 22-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcmasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-393-305-3

Query Match 15.4%; Score 99.5; DB 1; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00025;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

Qy 3 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCESAFSCF---QKAOLKSANTGNNE 53
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVTKAMKCFLELLEQLQVLSLESGDASIH 61
Qy 54 RIINVSIKKLRKPPSTNAGRQKRLTCPCSDSYEKPKPEFLERFKSLQKMIH 109
Db 62 TVENLII--LANNSSNGNVTES---GCKECEELEEKNIKEFLQSFVHVQMFN 112

RESULT 4

US-08-726-817-3
; Sequence 3, Application US/08726817
; Patent No. 5707616
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726.817

[illegible][illegible]

Query Match 15.4%; Score 99.5; DB 1; Length 114;
Best Local Similarity 25.2%; Pred. No. 0.00025;
Matches 30; Conservative 23; Mismatches 45; Indels 2

Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVISESGDXIHD 61
QY 53 --ERLINVSIKKLKKPPSTNAGRQKRLHLCPCSDSYEKPPKEFLERFKSLLOKMIH 109
Db 62 TVENLIILANLXSSNGNXTEG-----CKECELEEKNIKEFLQSFVHIVQMFN 112

RESULT 7

US-08-725-969-3
; Sequence 3, Application US/08725969
; Patent No. 5892001
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/725,969
; FILING DATE: 04-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,305
; FILING DATE: 22-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-725-969-3

Query Match 15.4%; Score 99.5; DB 2; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00025;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;
QY 3 IDVDQKKNVNDLVPEF-----LPAPEDVTNCWSAFSCF---QKAQLKSANTGNNE 53
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVISESGDASIH 61
QY 54 RIINVSIKKLKKPPSTNAGRQKRLHLCPCSDSYEKPPKEFLERFKSLLOKMIH 109
Db 62 TVENLIILANLXSSNGNXTEG-----CKECELEEKNIKEFLQSFVHIVQMFN 112

RESULT 8

US-08-794-524-3
; Sequence 3, Application US/08794524
; Patent No. 5985262
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk

; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/794,524
; FILING DATE: 03-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,305
; FILING DATE: 22-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-794-524-3

Query Match 15.4%; Score 99.5; DB 2; Length 114;
Best Local Similarity 25.9%; Pred. No. 0.00025;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;
QY 3 IDVDQKKNVNDLVPEF-----LPAPEDVTNCWSAFSCF---QKAQLKSANTGNNE 53
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVISESGDASIH 61
QY 54 RIINVSIKKLKKPPSTNAGRQKRLHLCPCSDSYEKPPKEFLERFKSLLOKMIH 109
Db 62 TVENLIILANLXSSNGNXTEG-----CKECELEEKNIKEFLQSFVHIVQMFN 112

RESULT 9

US-09-189-193-3
; Sequence 3, Application US/09189193
; Patent No. 6184359
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

RESULT 11
DCT-US94-03793-12

:
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Immunex Corporation
: STREET: 51 University Street
: CITY: Seattle
:

```

Query Match      15.4%; Score 99.5; DB 5; Length 114;
Best Local Similarity 25.2%; Pred. No. 0.00025;
Matches 30; Conservative 23; Mismatches 45; Indels 21; Gaps 5;

QY 3 IDVLDLKNVNDLVPFE-----LPAPEDYETNCESAFSCF-QKALQKLSANTGN----- 52
      :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 3 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCKVATMKCFLELQLVISEGDXIHD 61

OV 53 --ERITNVSTIKKLRKPPSTNAGRRKHKRTCTPSCSNYEKKPKPEETLERPKSLIOMKH 109

```


Qy	3	IDIVOLKKNYVNDLPEF-----LPAPEDVETNCESAFSCF-----QKAOLKSAANTGNNE	53
Db	51	VNVISDJKK-IEDDIQSMHIDATLYTSDVHPCKFTAMKCFLELLEQVLSLESGDAIHD	109
Qy	54	RIINVSIKKIKRRPPTNAGRRQKRLRITCPSCDSEYKKKKPKFELEFRKSLLOQMIIH	109
Db	110	TVENLIIIT--LANNSSLNGNVITES--GCKECELEBKNIKELQSFVHVIVQMFIN	160

Search completed: May 23, 2001, 11:09:42
Job time: 65 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:12:51 ; Search time 70.54 Seconds
(without alignments)
118.857 Million cell updates/sec

Title: US-09-522-217-2_COPY_41_162

Perfect score: 647

Sequence: 1 QLDIVDLKKNVNDLVPEF.....LLQKMIHOLLSRTHGSEDS 122

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_67:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	94.5	14.6	162	1 A53484	interleukin-15 pre
2	91.5	14.1	567	2 S29498	lymphocyte antigen
3	82	12.7	607	1 ABXL72	74K albumin precu
4	79	12.2	206	2 S49882	hypothetical prote
5	75.5	11.7	336	2 S42632	Fit-1S protein pre
6	74	11.4	2197	2 B71600	variant-specific s
7	73.5	11.4	162	2 I49124	interleukin-15 - m
8	73.5	11.4	868	2 T31527	hypothetical prote
9	73	11.3	304	2 A32108	translation initia
10	73	11.3	848	2 T00372	hypothetical prote
11	72.5	11.2	1101	2 T26919	hypothetical prote
12	72	11.1	262	2 F72858	probable methyl tr
13	72	11.1	262	2 T41813	ACMPV orf69 - Bom
14	71.5	11.1	195	2 S42022	ureidoglycolate hy
15	71	11.0	406	2 T28957	hypothetical prote
16	71	11.0	805	2 A46266	aryl hydrocarbon r
17	71	11.0	864	2 T08575	protein kinase hom
18	71	11.0	2206	2 G71611	hypothetical prote
19	71	11.0	2924	2 T19378	variant-specific s
20	70.5	10.9	309	2 T41889	PE38 orf153 - Bomb
21	70.5	10.9	447	2 S39316	CAB3b protein - hu
22	70.5	10.9	477	2 S21049	calcium channel pr
23	70.5	10.9	482	2 S41211	voltage-dependent
24	70.5	10.9	484	2 A46608	voltage-dependent
25	70.5	10.9	484	2 S39315	CAB3a protein - hu
26	70	10.8	405	2 A72383	sensor histidine k
27	70	10.8	1206	2 S24407	formin isoform IV
28	69.5	10.7	1468	2 S11515	formin - mouse
29	69.5	10.7	2129	2 T27431	hypothetical prote

30	69	10.7	1220	2 T06403	resistance complex
31	68.5	10.6	484	2 S62185	calcium channel be
32	68.5	10.6	491	2 JE0276	voltage-gated pota
33	68.5	10.6	1257	2 T01020	hypothetical prote
34	68.5	10.6	1671	2 S71628	sensory transducti
35	68	10.5	329	2 C69483	hypothetical prote
36	68	10.5	397	2 A33880	syndecan 2 - human
37	68	10.5	425	2 T50355	hypothetical prote
38	68	10.5	465	2 A56679	probable RNA-bindi
39	68	10.5	864	2 S60441	hypothetical prote
40	67.5	10.4	380	2 T25454	hypothetical prote
41	67.5	10.4	500	2 T28158	hypothetical prote
42	67.5	10.4	736	2 T00023	transcription fact
43	67.5	10.4	789	2 S44759	C14B9.5 protein -
44	67.5	10.4	1150	2 T13824	LK6-protein kinase
45	67.5	10.4	2734	2 B36579	inositol 1,4,5-tri

ALIGNMENTS

RESULT 1

A53484
interleukin-15 precursor - green monkey
C:Species: Cercopithecus aethiops (green monkey, grivet)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A53484
R:Grabstein, K.H.; Eisenman, J.; Shanebeck, K.; Rauch, C.; Srinivasan, S.; Fung, V.;
Giri, J.G.
Science 264, 965-968, 1994
A:Title: Cloning of a T cell growth factor that interacts with the beta chain of the
A:Reference number: A53484; MUID:94233380
A:Accession: A53484
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-162 <GRA>
A:Cross-references: GB:003099; NID:g493521; PIDN:AAA18416.1; PID:g493522
A:Note: the complete translation is not shown
C:Superfamily: interleukin-15
C:Keywords: growth factor
F:43-182/Product: interleukin-15 #status predicted <MAT>
F:83-133,90-136/Disulfide bonds: #status predicted

Query Match 14.6% Score 94.5; DB 1; Length 162;

Best Local Similarity 25.9%; Pred. No. 0.045; Mismatches 23; Indels 15; Gaps 5;

QY 3 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCESAFSCF----QRAQLKSANTGNNE 53

Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCKVTAMKCFELLEQVISHESGDTDHD 109

QY 54 RIINVSITKKLRKPPSTNAGRQKRLFCPCDSYSEKKPPKEFLERFKSLQKMH 109

Db 110 TVENLII--LANNTLSSNGNITES---GCKECELEEKNIKEFLQSFVHVQMFIN 160

RESULT 2

S29498
lymphocyte antigen Ly84 precursor - mouse
N:Alternate names: 38.5K T1 glycoprotein; ST2L protein
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Dec-2000
C:Accession: S29498; A33541; S17657; S07054
R:Yanagisawa, K.; Takagi, T.; Tsukamoto, T.; Tetsuka, T.; Tominaga, S.
FEBS Lett. 318, 83-87, 1993
A:Title: Presence of a novel primary response gene ST2L, encoding a product highly si
A:Reference number: S29498; MUID:93170492
A:Accession: S29498
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-567 <YAN>
A:Cross-references: EMBL:D13695; NID:g286100; PIDN:BAA02854.1; PID:g286101

Db 12 QATSVVNGLLSNLLPGVKIRANNGKTSVNNGSKAQLIDRLNKKRVQLQNRDVHKIKKCC 71
QY 68 PSTNAGRRQKHRLTCPCSDSYEKKKPPKEFLERF--KSLLOKMIHOHLSSRT 116
Db 72 KLVKKKKVKKHKL-----DKEQLQAKHQVLKK--HQHEGTLT 108

RESULT 5
S42632
Fit-15 protein precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 25-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 19-May-2000
C:Accession: S42632
R:Berger, G.; Reikerstorfer, A.; Braselmann, S.; Graninger, P.; Busslinger, M.
EMBO J. 13, 1176-1188, 1994
A:Title: Alternative promoter usage of the Fos-responsive gene Fit-1 generates mRNA isoforms
A:Reference number: S42632; MUID:94178260
A:Accession: S42632
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-336 <BER>
A:Cross-references: GB:U04319; NID:G488278; PIDN:AAA67172.1; PID:G488279
C:Superfamily: vaccinia virus B15R protein; immunoglobulin homology

Query Match 11.7%; Score 75.5; DB 2; Length 336;
Best Local Similarity 22.2%; Pred. No. 6.6;
Matches 30; Conservative 26; Mismatches 42; Indels 37; Gaps 6;

QY 3 IDIVDLKKNVNDLP-----EFLPAPEDVETNCWFAFSCFQKAKLQKSN 48
Db 49 INPVEWYSTNTNRIPTQKRNFVSRDLKFLPAKYE-----DSGIYTCVIRSP-ESIK 102

QY 49 TGNNERIINVSIRKKRPP-----STNAGRRQKHRLTCPCSDSYEKKKPPKEFLE 98
Db 103 TGS-----LNTIYK---RPNCKIPDYMSTVDGSDKSKITCPTIALYNWTPAVQWFK 155

QY 99 RFKSLLOKMIHQHLS 113
Db 156 NCRALQGPFRFRAHMS 170

RESULT 6
B71600
variant-specific surface protein 1 homolog PFB1055c - malaria parasite (Plasmodium falciparum)
N:Alternate names: erythrocyte membrane binding protein 1 (EMP1)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C:Accession: B71600
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743
A:Accession: B71600
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2197 <GAR>
A:Cross-references: GB:AE001434; GB:AE001362; NID:G3845341; PIDN:AAC71996.1; PID:G384534
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB1055c

Query Match 11.4%; Score 74; DB 2; Length 2197;
Best Local Similarity 29.5%; Pred. No. 64;
Matches 28; Conservative 16; Mismatches 41; Indels 10; Gaps 4;

QY 9 LKNVNDLPVEFLPAPEDVETNCWFAF--SCFQKAKLQKSNANTGNRIINVSIRKKR 66
Db 1594 VKSFLETWIPK-TAVNDQNVNKLKSGFGNCCGCSAISTN-GNEEDAIDCMIKKLEKK 1651

QY 67 -----PPSTNAGRRQKHRLTCPCSDSYEKKKPPKE 95

Db 1652 IDECRKPGENSGQTCNETLTHPLDVODEPLEE 1686
RESULT 7
I49124
interleukin-15 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 15-Mar-1996 #sequence_revision 15-Mar-1996 #text_change 16-Jul-1999
C:Accession: I49124
R:Anderson, D.M.; Johnson, L.; Glaccum, M.B.; Copeland, N.G.; Gilbert, D.J.; Jenkins, J.
Genomics 25, 701-706, 1995
A:Title: Chromosomal assignment and genomic structure of IL15.
A:Reference number: A56005; MUID:95278940
A:Accession: I49124
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-162 <RES>
A:Cross-references: EMBL:U14332; NID:G984941; PIDN:AAA75377.1; PID:G984942
C:Genetics:
A:Gene: IL15
C:Superfamily: interleukin-15

Query Match 11.4%; Score 73.5; DB 2; Length 162;
Best Local Similarity 19.8%; Pred. No. 4.8;
Matches 22; Conservative 30; Mismatches 54; Indels 5; Gaps 3;

QY 3 IDI---VQDKKNVNDL-VPEFLPAPEDVETNCWFAFSCFQKAKLQKSNANTGNRIIN 58
Db 51 IDVRYDLEKIESLIOSIHIDTTLTDSDFHPSCVKVAMNCF-LLEQLVILHEYSNMTLNE 109

QY 59 SIKLKKRPPSNAGRRQKHRLTCPCSDSYEKKKPPKEFLERPKSLLOKMIH 109
Db 110 TVRNVLYLANSTLSSNKNVAESGCELEBKTTFEFLQSFIRIVOMFIN 160

RESULT 8
T31527
hypothetical protein Y47D3A.14 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T31527
R:Matthews, L.
submitted to the EMBL Data Library, October 1999
A:Reference number: Z21043
A:Accession: T31527
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-868 <WIL>
A:Cross-references: EMBL:AL117202; PIDN:CAB55073.1; CESP:Y47D3A.14
A:Experimental source: clone Y47D3A
C:Genetics:
A:Gene: CESP:Y47D3A.14
A:Introns: 17/2; 51/3; 102/3; 169/3; 246/1; 287/2; 319/3; 371/3; 395/2; 440/3; 450/3;

Query Match 11.4%; Score 73.5; DB 2; Length 868;
Best Local Similarity 29.2%; Pred. No. 27;
Matches 19; Conservative 11; Mismatches 30; Indels 5; Gaps 1;

QY 33 WSFSCFQKAKLKS-----ANTGNRIINVSIRKKRPPSTNAGRRQKHRLTCPCSDS 87
Db 680 WSRVQAFRRKKNKSGGGTNSDSSEQSEKVLKLLKARRPSDETIVPPHYIVCPLFPS 739

QY 88 YEKKP 92
Db 740 HVAIP 744

RESULT 9
A32108
translation initiation factor eIF-2 alpha chain - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein J1429; protein YJR007w
C:Species: Saccharomyces cerevisiae
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 26-Aug-1999
C:Accession: A32108; S55195; S57022
R:Cigan, A.M.; Pabich, E.K.; Feng, L.; Donahue, T.F.
Proc. Natl. Acad. Sci. U.S.A. 86, 2784-2788, 1989
A:Title: Yeast translation initiation suppressor su12 encodes the alpha-subunit of eukaryotic translation initiation factor eIF-2
A:Reference number: A32108; MUID:89202411
A:Accession: A32108
A:Molecule type: DNA
A:Residues: 1-304 <CIG>
A:Cross-references: EMBL:M25552; NID:g341369; PIDN:AAA70332.1; PID:g903889
R:de Haan, M.; Smits, P.H.M.; Grivell, L.A.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55183
A:Accession: S55195
A:Molecule type: DNA
A:Residues: 1-304 <DEH>
A:Cross-references: EMBL:X87611; NID:g854567; PIDN:CAA60929.1; PID:g854580
R:de Haan, M.; Grivell, L.A.; Smits, P.H.M.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56771
A:Accession: S57022
A:Molecule type: DNA
A:Residues: 1-304 <ZAG>
A:Cross-references: EMBL:Z49507; NID:gl015631; PIDN:CAA89529.1; PID:gl015632; MIPS:YJR007w
C:Genetics:
A:Gene: SGD:SUI2
A:Cross-references: SGD:S0003767; MIPS:YJR007w
A:Map position: 10R
C:Superfamily: translation initiation factor eIF-2 alpha chain
C:Keywords: phosphoprotein; protein biosynthesis

Query Match 11.3%; Score 73; DB 2; Length 304;
Best Local Similarity 24.8%; Pred. No. 10;
Matches 34; Conservative 25; Mismatches 56; Indels 22; Gaps 7;

QY 4 DIVDQKNVND-LVPEFLPAPEDVETNC-EWSAFSCFOKAOLKSANTGNERRI-INVS1 60
Db 164 DVLDELKNYISKRLTPQAVKIRADVSCFSVEGIDAIDKA-LKSAEDMSTEQMVKVKL 222

QY 61 KKLKRPSTNAGRQKRLTCTPCSDSYEK-----KPPKEFLERFKSLIQ 105
Db 223 VAAPLYVLTQALDKQKIEQLS-AIEKITEVITYKGVGNITMPPRKAVTATEDAEIQ 280

QY 106 KMIH-OHLSSRTHGSED 121
Db 281 ALLESKELNRSDEDD 297

RESULT 10
T00372
hypothetical protein KIAA0650 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 17-Nov-2000
C:Accession: T00372; T12523
R:Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete coding sequence of the cDNA for KIAA0650, a novel human gene.
A:Reference number: T12523; MUID:98403880
A:Accession: T00372
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-848 <ISH>
A:Cross-references: EMBL:AB014550; NID:g3327113; PIDN:BA31625.1; PID:g3327114
R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A:Reference number: T12524
A:Accession: T12523
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 'E', 122-848 <WAM>
A:Cross-references: EMBL:AL080138
A:Experimental source: adult testis; clone DKF2p434K063
C:Genetics:
A:Note: KIAA0650; DKF2p434K063.1

Query Match 11.3%; Score 73; DB 2; Length 848;
Best Local Similarity 25.5%; Pred. No. 30;
Matches 27; Conservative 22; Mismatches 47; Indels 10; Gaps 5;

QY 19 ERTFAPEDVETNC-EWSAFSCFOKAOLKSANTGNERRIINVSIIKKLRKPPSTNAGRQK 77
Db 66 KFTPGPGNKDLCTWREFSDFIRVQLIS--GPPAKLLIDWPELKESIPVIN-GRDLQ 121

QY 78 HRLTCTPCSDSYEKKPPKEFLERFKSLQKMTQHL--SSRTHGSED 121
Db 122 NPIIVQLCDOWNPAP--VQHVKSITKASNLKLMPSNQOHTDE 164

RESULT 11
T26919
hypothetical protein Y45F10B.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26919
R:McMurray, A.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z20286
A:Accession: T26919
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1101 <WIL>
A:Cross-references: EMBL:AL021487; PIDN:CAA16357.1; GSPDB:GN00022; CESP:Y45F10B.10
A:Experimental source: clone Y45F10B
C:Genetics:
A:Gene: CESP:Y45F10B.10
A:Map position: 4
A:Introns: 217/3; 292/1; 451/3; 556/2; 637/3; 755/3; 823/2; 946/1; 1078/3

Query Match 11.2%; Score 72.5; DB 2; Length 1101;
Best Local Similarity 31.5%; Pred. No. 44;
Matches 23; Conservative 14; Mismatches 23; Indels 13; Gaps 4;

QY 40 OKAOLKSANTGNERRIINVSIIKKLRKP--PSTNAGRQK-HRLTCTPCSDSYEKKPPKEF 96
Db 1013 RSARQSVSSASNEPVASTSAGEIKDPILSSNNGGNAQSAPRATAP-----KPTFDM 1065

QY 97 LERFK---SLLOK 106
Db 1066 LERSKRTSLIEK 1078

RESULT 12
F72858
probable methyl transferase - Autographa californica nuclear polyhedrosis virus
C:Species: Autographa californica nuclear polyhedrosis virus, ACNPNV
A:Note: dsDNA virus
C:Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 12-Nov-1999
C:Accession: F72858
R:Ayres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.
Virology 202, 586-605, 1994
A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis vir
A:Reference number: A72850; MUID:94303173
A:Accession: F72858
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <AYR>
A:Cross-references: GB:L22858; NID:g510708; PIDN:AAA66699.1; PID:g5559138
C:Genetics:
A:Gene: AcOrf-69

```

Query Match      11.18; Score 72; DB 2; Length 262;
Best Local Similarity 21.7%; Pred. No. 11;
Matches 18; Conservative 20; Mismatches 35; Indels 10; Gaps 2;

QY 30 NCESAFSCFOKAOLKSANTGNERIINVSTIKLKRKPPSTNAGRROKH-----RLTCPS 84
Db 176 NCVLKVFDAFHEHTIOMLN-----KFVNHFKEWVLYKPPSSRPANSERYLICFNKLVRPY 230
    ||| | : : : | : : : | : : | : | : | : | : | : | : | : | : |
QY 85 CDSYEKKPKPEFLERFKSLLOKM 107
Db 231 CNNVNELEKQFEKYRIQLKNL 253
    |:| : |:| : |:| : |:| : |:| : |:| : |:| : |:| : |:| : |:| :

RESULT 13
T41813
ACMPNP orf69 - Bombyx mori nuclear polyhedrosis virus (isolate T3)
C:Species: Bombyx mori nuclear polyhedrosis virus, BmsNPV
A:Variety: isolate T3
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C:Accession: T41813
R:Gomi, S.; Majima, K.; Maeda, S.
J. Gen. Virol. 80, 1323-1337, 1999
A>Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
A:Reference number: Z22020; MUID:99281911
A:Accession: T41813
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-262 <KAM>
A:Cross-references: EMBL:L33180; PIDN:AAC63742.1
A:Experimental source: isolate T3
C:Genetics:
A>Note: Orf_57

Query Match      11.18; Score 72; DB 2; Length 262;
Best Local Similarity 21.7%; Pred. No. 11;
Matches 18; Conservative 19; Mismatches 36; Indels 10; Gaps 2;

QY 30 NCESAFSCFOKAOLKSANTGNERIINVSTIKLKRKPPSTNAGRROKH-----RLTCPS 84
Db 176 NCVLKVFDAFHEHTIOMLN-----KFVNHFKEWVLYKPPSSRPANSERYLICFNKLVRPY 230
    ||| | : : : | : : : | : : | : | : | : | : | : | : | : | : |
QY 85 CDSYEKKPKPEFLERFKSLLOKM 107
Db 231 CNNVNELEKQFEKYRIQLKNL 253
    |:| : |:| : |:| : |:| : |:| : |:| : |:| : |:| : |:| : |:| :

RESULT 14
S42022
ureidoglycolate hydrolase (EC 3.5.3.19) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YIR032c
C:Species: Saccharomyces cerevisiae
C:Date: 31-Mar-1992 #sequence_revision 14-Sep-1994 #text_change 21-Jul-2000
C:Accession: S42022; S48494
R:Yoo, H.S.; Cooper, T.G.
Yeast 7, 693-698, 1991
A>Title: The ureidoglycolate hydrolase (DAL3) gene in Saccharomyces cerevisiae.
A:Reference number: S42022; MUID:92133160
A:Accession: S42022
A:Molecule type: DNA
A:Residues: 1-195 <YOO>
A:Cross-references: EMBL:M64778; NID:g171369; PIDN:AAA73025.1; PID:g171370
R:Rowley, K.
submitted to the EMBL Data Library, October 1994
A:Reference number: S48478
A:Accession: S48494
A:Molecule type: DNA
A:Residues: 1-195 <ROW>
A:Cross-references: GB:247047; EMBL:Z38061; NID:g603997; PID:g763377; MIPS:YIR032c
C:Genetics:
A:Gene: SGD:DAL3
A:Cross-references: SGD:S0001471; MIPS:YIR032c

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:20:24 ; Search time 40.06 Seconds
(without alignments)
104.323 Million cell updates/sec

Title: US-09-522-217-2_COPY_41_162

Perfect score: 647

Sequence: 1 QLIDIVDLKNVNDLVPEF.....LLQKMIHQHSRTHGSEDS 122

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101.5	15.7	162	1 IL15_FELCA	O97687 felis silve
2	99.5	15.4	162	1 IL15_HUMAN	P40933 homo sapien
3	94.5	14.6	162	1 IL15_CERAE	P40221 cercopithe
4	94.5	14.6	162	1 IL15_MACMU	P48092 macaca mula
5	91.5	14.1	337	1 IRL1_MOUSE	P14719 mus musculu
6	90.5	14.0	162	1 IL15_BOVIN	Q28028 bos taurus
7	82	12.7	607	1 ALB2_XENLA	P14872 xenopus lae
8	80.5	12.4	162	1 IL15_PIG	Q95253 sus scrofa
9	79	12.2	206	1 IIM7_YEAST	P40470 saccharomyc
10	73.5	11.4	162	1 IL15_MOUSE	P48346 mus musculu
11	73	11.3	304	1 IF2A_YEAST	P20459 saccharomyc
12	72	11.1	262	1 Y069_NPVAC	P41469 autographa
13	71.5	11.1	195	1 ALIA_YEAST	P32459 saccharomyc
14	71.5	11.1	484	1 CCB3_MOUSE	P54285 mus musculu
15	71	11.0	805	1 AHR_MOUSE	P30561 mus musculu
16	70.5	10.9	477	1 CCB3_RABIT	P54286 oryctolagus
17	70.5	10.9	484	1 CCB3_HUMAN	P54284 homo sapien
18	70.5	10.9	484	1 CCB3_RAT	P54287 rattus norv
19	70	10.8	1206	1 FM14_MOUSE	Q05859 mus musculu
20	70	10.8	1222	1 YMP3_CAEEL	Q10947 caenorhabdi
21	69.5	10.7	1468	1 FMN1_MOUSE	Q05860 mus musculu
22	69	10.7	741	1 IDH_AZOV1	P16100 azotobacter
23	68.5	10.6	162	1 IL15_RAT	P97604 rattus norv
24	68	10.5	132	1 IL4_CANFA	O77762 canis fami
25	68	10.5	201	1 SDG2_HUMAN	P34741 homo sapien
26	68	10.5	864	1 IG3M_YEAST	P48237 saccharomyc
27	67.5	10.4	300	1 SPY4_MOUSE	O9wtp2 mus musculu
28	67.5	10.4	464	1 N2B_HAEIR	P46441 haematobia
29	67.5	10.4	736	1 BAC1_HUMAN	O14867 homo sapien
30	67.5	10.4	789	1 YK25_CAEEL	P34332 caenorhabdi
31	67.5	10.4	2749	1 IP3R_RAT	P29994 rattus norv
32	67	10.4	2476	1 ATRX_MOUSE	O61687 mus musculu
33	66.5	10.3	420	1 VIT3_DROME	P06607 drosophila

34	66.5	10.3	441	1 FUS6_ARATH	P45432 arabidopsis
35	66.5	10.3	739	1 BAC1_MOUSE	P97302 mus musculu
36	66.5	10.3	1750	1 Y832_METJA	O58242 methanococc
37	66.5	10.3	3712	1 ACVS_CEPAC	P25464 cephalospor
38	66	10.2	345	1 TF2B_YEAST	P29055 saccharomyc
39	66	10.2	822	1 MUS2_SYNY3	P73625 synchocyst
40	66	10.2	932	1 PMS1_HUMAN	P54277 homo sapien
41	66	10.2	2375	1 ATRX_HUMAN	P46100 homo sapien
42	65.5	10.1	525	1 KBN8_YEAST	P38070 saccharomyc
43	65.5	10.1	562	1 BS69_HUMAN	Q15326 homo sapien
44	65.5	10.1	576	1 P80C_HUMAN	P38432 homo sapien
45	65	10.0	472	1 IFT2_HUMAN	P09913 homo sapien

ALIGNMENTS

RESULT 1

IL15_FELCA

ID IL15_FELCA STANDARD; PRT: 162 AA.

AC O97687;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE INTERLEUKIN-15 PRECURSOR (IL-15).

GN IL15.

OS Felis silvestris catus (Cat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

OX NCBI_TaxID=9685;

RP [1]

RN SEQUENCE FROM N.A.

RC TISSUE=Lymph node;

RA Barger A.B., Dean G.A., Lavoy A.S.;

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

CC Lymphocytes. STIMULATION BY IL-15 REQUIRES INTERACTION OF T-

CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R

CC GAMMA BUT NOT IL-2R ALPHA (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: SECRETED.

CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL; AF108148; AAD05268.1; -

KW Cytokine; Glycoprotein; Signal.

FT SIGNAL 1 29 POTENTIAL.

FT PROPEP 30 48 POTENTIAL.

FT CHAIN 49 162 INTERLEUKIN-15.

FT DISULFID 83 133 POTENTIAL.

FT DISULFID 90 136 POTENTIAL.

FT CARBOHYD 104 104 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 127 127 N-LINKED (GLCNAC...) (POTENTIAL).

SQ SEQUENCE 162 AA; 18412 MW; D8C7CE7F40110DD CRC64;

Query Match 15.7%; Score 101.5; DB 1; Length 162;

Best Local Similarity 26.5%; Pred. No. 0.0047;

Matches 31; Conservative 19; Mismatches 48; Indels 19; Gaps 5;

Qy 4 DIVDQLK---NYVNDL-VPEFLPAPEDVTNCEWSAFSCF----QKAQLKSANTGNRI 55

Db 52 DVISDLKIIDKTIQSLHIDATLYTSDVHPNCKVPTAMKCFLLLEHLHVLSLEKNETIHQT 111

Qy 56 INVSTKKLKRPPSPNAGRRQKRLT---CPSCDSYEKKPKPKFLERFKSLQKMIH 109

Db 112 ENIII-----LANSGLSSNRNITETCKCEELEKNEIKFLOFSFHVHVMFIN 160

RESULT 2
IL15_HUMAN STANDARD; PRT; 162 AA.
ID IL15_HUMAN
AC P40933; O93058; O43512;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN RP SEQUENCE FROM N.A. (IL15-S48AA ISOFORM).
RC TISSUE=Bone marrow;
RX MEDLINE=94233380; PubMed=8178155;
RA Grabstein K.K., Eisenman J., Shanebeck K., Rauch C.,
RA Srinivasan S., Fung V., Beers C., Richardson J., Schoenborn M.A.,
RA Adhieh M., Johnson L., Alderson M.R., Watson J.D., Anderson D.M.,
RA Giri J.G.;
RT "Cloning of a T cell growth factor that interacts with the beta chain
of the interleukin-2 receptor.";
RL Science 264:965-968(1994).
[2]
RN RP SEQUENCE FROM N.A. (IL15-S48AA ISOFORM).
RA Krause H., Jandrig B., Wernicke C., Bullfong-Paus S., Pohl T.,
RA Diamantstein T.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
[3]
RN RP SEQUENCE FROM N.A. (IL15-S21AA ISOFORM).
RC TISSUE=Lung cancer;
RX MEDLINE=96218668; PubMed=8668345;
RA Meazza R., Verdiani S., Biassoni R., Coppolecchia M., Gaggero A.,
RA Orengo A.M., Colombo M.P., Azzone B., Ferrini S.;
RT Identification of a novel interleukin-15 (IL-15) transcript isoform
generated by alternative splicing in human small cell lung cancer
cell lines.";
RL Oncogene 12:2187-2192(1996).
[4]
RN RP SEQUENCE FROM N.A. (IL15-S21AA ISOFORM).
RC TISSUE=Testis;
RX MEDLINE=98070771; PubMed=9405632;
RA Tadaya Y., Kurys G., Thies T.A., Losi J.M., Azimi N., Hanover J.A.,
RA Bamford R.N., Waldmann T.A.;
RT "Generation of secretable and nonsecretable interleukin 15 isoforms
through alternate usage of signal peptides.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:14444-14449(1997).
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
GAMMA BUT NOT IL-2R ALPHA.
CC -!- SUBCELLULAR LOCATION: SECRETED (IL15-S48AA). IL15-S21AA IS NOT
SECRETED, BUT RATHER IS STORED INTRACELLULARLY, APPEARING IN THE
NUCLEUS AND CYTOPLASMIC COMPONENTS.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; IL15-S48AA (SHOWN HERE) AND
IL15-S21AA; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: MOST ABUNDANT IN PLACENTA AND SKELETAL MUSCLE.
CC IT IS ALSO DETECTED IN THE HEART, LUNG, LIVER AND KIDNEY. IL15-
S21AA IS PREFERENTIALLY EXPRESSED IN TISSUES SUCH AS TESTIS AND
THYMUS.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC -!- DATABASE: NAME=R&D Systems' cytokine source book;
CC WWW="http://www.rndsystems.com/cyt_cat/il-15.html".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
or send an email to license@isb-sib.ch.

CC or send an email to license@isb-sib.ch.
CC
DR EMBL; U14407; AA21551.1; -
DR EMBL; X91233; CAAG2616.1; -
DR EMBL; X94223; CAAG3914.1; -
DR EMBL; X94222; CAAG3913.1; -
DR EMBL; AF031167; AAB97518.1; -
DR MIM; 600554; -
KW Cytokine; Glycoprotein; Signal; Alternative splicing.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48 POTENTIAL.
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1 37 MRISKPHLSISITOCYCLLLNSHFLTEAGIRHVFILG ->
FT MVLTIDILCS (IN ISOFORM IL15-S21AA).
FT CONFLICT 141 141 E -> K (IN REF. 4).
SQ SEQUENCE 162 AA; 18086 MW; 0CE0520C1D8379E2 CRC64;

Query Match 15.4%; Score 99.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.0073;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 3 IDVDOLKNYVNDLVPEF-----LPAPEDVETNCESAFSCF-----OKAOLKSANTGNNE 53
Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPSCVKVTAMKCFLELQVLSLESGDASHD 109
QY 54 RIINVSIKLKRPPSTNAGRROKHRLTCPCSDSYEKPKPKPELERFKSLQKMIH 109
Db 110 TVENLI--LANNLSNGNVTES---GCKEKELEEKNIKEFLQSEFVHVQMFIN 160

RESULT 3
IL15_CERAE STANDARD; PRT; 162 AA.
ID IL15_CERAE
AC P40221;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 49-81.
RC TISSUE=Kidney;
RX MEDLINE=94233380; PubMed=8178155;
RA Grabstein K.K., Eisenman J., Shanebeck K., Rauch C.,
RA Srinivasan S., Fung V., Beers C., Richardson J., Schoenborn M.A.,
RA Adhieh M., Johnson L., Alderson M.R., Watson J.D., Anderson D.M.,
RA Giri J.G.;
RT "Cloning of a T cell growth factor that interacts with the beta chain
of the interleukin-2 receptor.";
RL Science 264:965-968(1994).
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
GAMMA BUT NOT IL-2R ALPHA.
CC -!- SUBCELLULAR LOCATION: SECRETED
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
or send an email to license@isb-sib.ch.

```
CC -----
DR EMBL: U03099; AAA18416.1; -.
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 127 127 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 162 AA; 18222 MW; 1BF9A82644E1C9B7 CRC64;

Query Match 14.6%; Score 94.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.022;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

OY 3 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCSEAFSCF-----OKAQLKSANTGNNE 53
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTAMKCFLEELQVISHESGDTDIHD 109
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 54 RIINVSITKKLRPPSTNAGRQKRLTQPCSDSYEKPKPKPEFLERFKSLQKMIH 109
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 110 TVENLII--LANNILSSNGNITES---GCKECELEEKNIKEFLQSFVHIVQMFN 160
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
IL15_MACMU
ID IL15_MACMU STANDARD; PRT; 162 AA.
AC P48092;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=96003435; PubMed=7561102;
RA Willinger F.J., Brar S.S., Wayne A.E., Chikkala N., Ansari A.A.;
RT "Comparative sequence analysis of cytokine genes from human and
RL nonhuman primates.";
RN J. Immunol. 155:3946-3954(1995).
RP SEQUENCE FROM N.A.
RA Tatsumi M.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U19843; AAB60398.1; -.
DR EMBL: AB000555; BAA19149.1; -.
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
```

```
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 127 127 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARIANT 5 5 K -> T.
FT VARIANT 31 31 I -> T.
SQ SEQUENCE 162 AA; 18194 MW; D233CF7FF6188C01 CRC64;

Query Match 14.6%; Score 94.5; DB 1; Length 162;
Best Local Similarity 25.9%; Pred. No. 0.022;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

OY 3 IDIVDLKKNVNDLVPEF-----LPAPEDVETNCSEAFSCF-----OKAQLKSANTGNNE 53
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 51 VNVISDLKK-IEDLIQSMHIDATLYTESDVHPCKVTAMKCFLEELQVISHESGDTDIHD 109
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 54 RIINVSITKKLRPPSTNAGRQKRLTQPCSDSYEKPKPKPEFLERFKSLQKMIH 109
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 110 TVENLII--LANNILSSNGNITES---GCKECELEEKNIKEFLQSFVHIVQMFN 160
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
IRL1_MOUSE
ID IRL1_MOUSE STANDARD; PRT; 337 AA.
AC P14719;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE INTERLEUKIN 1 RECEPTOR-LIKE 1 PRECURSOR (ST2 PROTEIN) (T1 PROTEIN)
DE (LYMPHOCYTE ANTIGEN 84).
GN IL1RL1 OR ST2 OR ST2 OR LY84.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=90092495; PubMed=2532153;
RA Tominaga S.;
RT "A putative protein of a growth specific cDNA from BALB/c-3T3 cells
RL is highly similar to the extracellular portion of mouse interleukin 1
RN receptor.";
RL FEBS Lett. 258:301-304(1989).
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/HE; TISSUE=Spleen;
RX MEDLINE=91355215; PubMed=1832015;
RA Tominaga S.I., Jenkins N.A., Gilbert D.J., Copeland N.G.,
RA Tetsuka T.;
RT "Molecular cloning of the murine ST2 gene. Characterization and
RL chromosomal mapping.";
RL Biochim. Biophys. Acta 1090:1-8(1991).
RP SEQUENCE FROM N.A.
RX MEDLINE=89345536; PubMed=2527364;
RA Klemenz R., Hoffmann S., Werenskiold A.K.;
RT "Serum- and oncoprotein-mediated induction of a gene with sequence
RL similarity to the gene encoding carcinoembryonic antigen.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:5708-5712(1989).
CC -!- FUNCTION: POSSIBLY INVOLVED IN REGULATION OF T-LYMPHOCYTE
CC ACTION.
CC -!- DEVELOPMENTAL STAGE: GROWTH-SPECIFIC EXPRESSION, G1-PHASE OF
CC CELL CYCLE.
CC -!- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -!- SIMILARITY: STRONG, TO INTERLEUKIN-1 RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC EMBL: Y07519; CAA68812.1; -
DR EMBL: X60184; CAA42742.1; -
DR EMBL: M24843; AAA40160.1; -
DR PIR: S07054; S07054.
DR MGI: 98427; Ly84.
DR InterPro: IPR003006; -
DR Pfam: PF00047; Ig; 3.
KW Immunoglobulin domain; Glycoprotein; Signal.
FT SIGNAL 1 26
FT CHAIN 27 337
FT DOMAIN 35 100
FT DOMAIN 132 194
FT DOMAIN 233 315
FT DISULFID 42 93
FT DISULFID 139 187
FT DISULFID 240 308
FT CARBOHYD 60 60
FT CARBOHYD 101 101
FT CARBOHYD 107 107
FT CARBOHYD 146 146
FT CARBOHYD 176 176
FT CARBOHYD 194 194
FT CARBOHYD 225 225
FT CARBOHYD 259 259
FT CARBOHYD 278 278
FT VARIANT 192 192
FT SEQUENCE 337 AA; 38502 MW; 7574372722486926 CRC64;
SQ

Query Match 14.18; Score 91.5; DB 1; Length 337;
Best Local Similarity 25.68; Pred. No. 0.096;
Matches 33; Conservative 20; Mismatches 39; Indels 37; Gaps 7;

Qy 7 DQLKNYNDVPEFLPAPEDVETNCESAFSCFQAKLKSANTGNINYSIKLKRK 66
Db 76 DRLK-----FLPQVE-----DSGIYACV-----IRSPNL-NKTYLVNTIHK---K 113

Qy 67 PPSTN-----AGRRQKRLTCPCSDSYEKPKPKFELFKSLQKMHQH----- 111
Db 114 PPSNIPDLYMYSTVRGSDKNFKITCPTIDLYNWTAPVQWFKNCALQEPFRHRSYLF 173

Qy 112 LSSRTHGSE 120
Db 174 IDNVTHDDE 182

RESULT 6
IL15_BOVIN
ID IL15_BOVIN STANDARD; PRT; 162 AA.
AC Q28028;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HOLSTEIN;
RX MEDLINE=97426124; PubMed=9282828;
RA Canals A., Gasbarre L.C., Boyd P.C., Almeria S., Zarlenga D.S.;
RT "Cloning and expression of bovine interleukin-15: analysis and
RT modulation of transcription by exogenous stimulation."
RL J. Interferon Cytokine Res. 17:473-480(1997).
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA (BY SIMILARITY).

```

```

CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U42433; AA851130.1; -
DR Cytokine; Glycoprotein; Signal.
KW SIGNAL 1 29
FT PROPEP 30 48
FT CHAIN 49 162
FT DISULFID 83 133
FT DISULFID 90 136
FT CARBOHYD 104 104
FT CARBOHYD 113 113
FT CARBOHYD 121 121
FT CARBOHYD 127 127
FT SEQUENCE 162 AA; 18554 MW; 6A63CAA329EB8302 CRC64;
SQ

Query Match 14.08; Score 90.5; DB 1; Length 162;
Best Local Similarity 25.5%; Pred. No. 0.053;
Matches 28; Conservative 17; Mismatches 58; Indels 7; Gaps 3;

Qy 5 IVDLKNYNDVPEF-----LPAPEDVETNCESAFSCFQAKLKSANTGNININVS 59
Db 53 VINDLKT-IEHLQSIHMDATLYTESDAHPNCKVMTAQCFLELRLVILHESKNATYIEI- 110

Qy 60 IKKLKRPSTNAGRKHLRITCPCSDSYEKPKPKFELFKSLQKMH 109
Db 111 IENLTMLANSLSIENKTELGCCEKELEKSIKELKSFVHVQMFN 160

RESULT 7
ALB2_XENLA
ID ALB2_XENLA STANDARD; PRT; 607 AA.
AC P14872;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE 01-NOV-1995 (Rel. 32, Last annotation update)
DE 74 KDA SERUM ALBUMIN PRECURSOR.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE OF 3-607 FROM N.A.
RX MEDLINE=89313788; PubMed=2747653;
RA Moskaitis J.E., Sargent T.D., Smith L.H. Jr., Pastori R.L.,
RA Schoenberg D.R.;
RT "Xenopus laevis serum albumin: sequence of the complementary
RT deoxyribonucleic acids encoding the 68- and 74-kilodalton peptides
RT and the regulation of albumin gene expression by thyroid hormone
RT during development."
RL Mol. Endocrinol. 3:464-473(1989).
RN [2]
RP SEQUENCE OF 1-48 FROM N.A.
RX MEDLINE=88172470; PubMed=2451026;
RA Schorpp M., Doebbeling U., Wagner U., Ryffel G.U.;
RT "5'-flanking and 5'-proximal exon regions of the two xenopus albumin
RT genes. Deletion analysis of constitutive promoter function."
RL J. Mol. Biol. 199:83-93(1988).
RN [3]
RP SEQUENCE OF 459-557 FROM N.A.
RX TISSUE=Liver;
RX MEDLINE=85126974; PubMed=3971963;
RA Wolffe A.P., Glover J.F., Martin S.C., Tenniswood M.P.R.,

```

RA Williams J.L., Tata J.R.;
 RT "Deinduction of transcription of Xenopus 74-kDa albumin genes and
 RT destabilization of mRNA by estrogen in vivo and in hepatocyte
 RT cultures";
 RL Eur. J. Biochem. 146:489-496(1985).
 CC -1- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD
 CC BINDING CAPACITY FOR WATER, CA⁺⁺, NA⁺, K⁺, FATTY ACIDS, HORMONES,
 CC BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
 CC COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M21442; AAA49637.1; -
 CC EMBL: M28276; AAA49642.1; -
 CC PIR: B41682; ABXL72.
 CC HSP: P02768; IUOR.
 CC InterPro: IPR000264; -
 CC Pfam: PF00273; transport_prot; 3.
 CC PRINTS: P00802; SERUMALBUMIN.
 CC PROSITE: PS00212; ALBUMIN; 3.
 CC Plasma; Metal-binding; Lipid-binding; Albumin; Repeat; Signal;
 CC Copper.
 KW SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 24 POTENTIAL.
 FT CHAIN 25 607 74 KDA SERUM ALBUMIN.
 FT REPEAT 29 211 1.
 FT REPEAT 217 403 2.
 FT REPEAT 409 601 3.
 FT METAL 30 30 COPPER (BY SIMILARITY).
 FT DISULFID 80 88 BY SIMILARITY.
 FT DISULFID 101 117 BY SIMILARITY.
 FT DISULFID 116 127 BY SIMILARITY.
 FT DISULFID 147 192 BY SIMILARITY.
 FT DISULFID 191 200 BY SIMILARITY.
 FT DISULFID 223 269 BY SIMILARITY.
 FT DISULFID 268 276 BY SIMILARITY.
 FT DISULFID 288 302 BY SIMILARITY.
 FT DISULFID 301 312 BY SIMILARITY.
 FT DISULFID 339 384 BY SIMILARITY.
 FT DISULFID 383 392 BY SIMILARITY.
 FT DISULFID 415 461 BY SIMILARITY.
 FT DISULFID 460 471 BY SIMILARITY.
 FT DISULFID 484 500 BY SIMILARITY.
 FT DISULFID 499 510 BY SIMILARITY.
 FT DISULFID 537 582 BY SIMILARITY.
 FT DISULFID 581 590 BY SIMILARITY.
 FT CONFLICT 503 503 S -> L (IN REF. 3).
 FT CONFLICT 531 531 H -> D (IN REF. 3).
 SQ SEQUENCE 607 AA; 70382 MW; 592BA4177A36B66B CRC64;
 Query Match 12.7%; Score 82; DB 1; Length 607;
 Best Local Similarity 24.2%; Pred. No. 1.5;
 Matches 38; Conservative 19; Mismatches 50; Indels 50; Gaps 6;
 QY 4 DIVDQLKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKAQLKSAN 48
 DB 72 EINFDAKSCINDICEKPVGTLFFDKLCADPAVGVNVEWSKECCAKQDPERACQCFKAH 131
 QY 49 TGNNERII---NVSIKKLKRK-----PPSTNAGRRQKHRLTCPS 84
 DB 132 RDHEHTSIRPEETCKLKEKHPDDLLSAFHEARNHPDLYPPAVLALTKQYHKLAEHC 191

QY 85 CDSYEKKPKPEFLERFKSLLOKMIHOHLSSRTHGSED 121
 DB 192 CEEDKE--KCFSEKMKQLMK-----QSHSIED 217
 RESULT 8
 IL15_PIG
 ID IL15_PIG STANDARD; PRT; 162 AA.
 AC Q95253;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE INTERLEUKIN-15 PRECURSOR (IL-15).
 GN IL15.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=97449311; PubMed=9305780;
 RA Canals A., Grimm D.R., Gasbarre L.C., Lunney J.K., Zarlenga D.S.;
 RT "Molecular cloning of cDNA encoding porcine interleukin-15";
 RL Gene 195:337-339(1997).
 CC -1- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
 CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
 CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
 CC GAMMA BUT NOT IL-2R ALPHA (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U58142; AAB72031.1; -
 KW Cytokine; Glycoprotein; Signal.
 FT SIGNAL 1 29 POTENTIAL.
 FT PROPEP 30 48 POTENTIAL.
 FT CHAIN 49 162 INTERLEUKIN-15.
 FT DISULFID 83 133 POTENTIAL.
 FT DISULFID 90 136 POTENTIAL.
 FT CARBOHYD 119 119 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 162 AA; 18437 MW; 7EF7992391883446 CRC64;
 Query Match 12.4%; Score 80.5; DB 1; Length 162;
 Best Local Similarity 24.2%; Pred. No. 0.48;
 Matches 29; Conservative 20; Mismatches 44; Indels 27; Gaps 6;
 QY 5 IVDQLKNVNDLVPEF-----LPAPEDVETNCWSAFSCF-----QKAQLKSANTG 50
 DB 53 VISDLKK-IEDLIRSIHMDATLYTESDAHPNCKVTAMKCFLLRLRVLQESRNSIDTV 111
 QY 51 NNERII-NVSIKKLKRKPPSTNAGRRQKHRLTCPSDSEYKPKPEFLERFKSLLOKMIH 109
 DB 112 ENLIILANSLSISIEYK---TESG-----CKECELEKKNINEFLKSFTIHVQMFIN 160
 RESULT 9
 YIM7_YEAST
 ID YIM7_YEAST STANDARD; PRT; 206 AA.
 AC P40470;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HYPOTHETICAL 23.8 KDA PROTEIN IN MET18-STH1 INTERGENIC REGION.
 GN YIL127C.

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churcher S.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gentles S., Hamlyn N., Hornselli T.S., Hunt S., Jagels K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Walsh S.V., Whitehead S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z46833; CAA86865.1; --
DR SGD; S0001389; YIL127C.
KW Hypothetical protein.
SQ SEQUENCE 206 AA; 23845 MW; 4E51B11216F7ECB8 CRC64;

Query Match 12.2%; Score 79; DB 1; Length 206;
Best Local Similarity 26.1%; Pred. No. 0.87;
Matches 29; Conservative 18; Mismatches 48; Indels 16; Gaps 3;

Oy 8 QLNKYNVDLPPEFLPAPEDVETNCSEAFSCFQKALKSANTGNNERINVSIIKKLRKP 67
Db 12 QATSVVGNLLSNLLPGVPIKIRANNGKSVNNGSKAQLIDRNKKRVQLQNRDQVHKKKC 71

Oy 68 PSTNAGRQKRLTCPCSDSYEKPKPEFLERF--KSLLQKMIHQLSSRT 116
Db 72 KLVKKKKVKKHKL-----DKQEQLEQAKRHVLKK--HQEGTGLT 108

RESULT 10
IL15_MOUSE
ID IL15_MOUSE STANDARD; PRT; 162 AA.
AC P48346;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WC/REJ X C57BL/6J; TISSUE=Bone marrow;
RX MEDLINE=95278940; PubMed=7759105;
RA Anderson D.M., Johnson L., Glaccum M.B., Copeland N.G.,
RA Gilbert D.J., Jenkins N.A., Valentine V., Kirstein M.N.,
RA Shapiro D.N., Morris S., Grabstein K., Cosman D.;
RT "Chromosomal assignment and genomic structure of IL15.";
RL Genomics 25:701-706(1995).
CC
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA.
CC
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U14332; AAA75377.1; --
DR MGD; MGI:103014; I115.
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48 POTENTIAL.
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT DISULFID 90 136 POTENTIAL.
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 108 108 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 162 AA; 18593 MW; 68C971498CEBF296 CRC64;

Query Match 11.4%; Score 73.5; DB 1; Length 162;
Best Local Similarity 19.8%; Pred. No. 2.2;
Matches 22; Conservative 30; Mismatches 54; Indels 5; Gaps 3;

Oy 3 IDI---VDQKNVYNDL-VPEFLPAPEDVETNCSEAFSCFQKALKSANTGNNERINV 58
Db 51 IDVRYDEKIESLIQSIHIDTTLTSDSFHPSCKVTAMNCF-LLEQLVILHEYSNMTLNE 109

Oy 59 SIKLKRKPPSTNAGRQKRLTCPCSDSYEKPKPEFLERFESLLQKMIH 109
Db 110 TVRNVLILANSTLSSKNVAESGCKEELKFTFEFLQSFIRIVQMFN 160

RESULT 11
IF2A_YEAST
ID IF2A_YEAST STANDARD; PRT; 304 AA.
AC P20459;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE EUKARYOTIC TRANSLATION INITIATION FACTOR 2 ALPHA SUBUNIT (EIF-2-
DE ALPHA).
GN SUI2 OR TIF211 OR YJR007W OR J1429.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89202411; PubMed=2649894;
RA Cigan A.M., Pabich E.K., Feng L., Donahue T.F.;
RT "Yeast translation initiation suppressor sui2 encodes the alpha
RT subunit of eukaryotic initiation factor 2 and shares sequence
RT identity with the human alpha subunit.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:2784-2788(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FYJ1679;
RA de Haan M., Smits P.H.M., Grivell L.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: EIF-2 FUNCTIONS IN THE EARLY STEPS OF PROTEIN SYNTHESIS
CC BY FORMING A TERNARY COMPLEX WITH GTP AND INITIATOR TRNA. THIS
CC COMPLEX BINDS TO A 40S RIBOSOMAL SUBUNIT, FOLLOWED BY MRNA BINDING
CC TO FORM A 43S PREINITIATION COMPLEX. JUNCTION OF THE 60S RIBOSOMAL
CC SUBUNIT TO FORM THE 80S INITIATION COMPLEX IS PRECEDED BY
CC HYDROLYSIS OF THE GTP BOUND TO EIF-2 AND RELEASE OF AN EIF-2-GDP
CC BINARY COMPLEX. IN ORDER FOR EIF-2 TO RECYCLE AND CATALYZE ANOTHER
CC ROUND OF INITIATION, THE GDP BOUND TO EIF-2 MUST EXCHANGE WITH GTP
CC BY WAY OF A REACTION CATALYZED BY EIF-2B.
CC
CC -!- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA, AND A GAMMA
CC CHAIN.
CC
CC -!- SIMILARITY: BELONGS TO THE EIF-2 ALPHA FAMILY.
CC
CC -!- SIMILARITY: CONTAINS 1 'S1 MOTIF' DOMAIN.

QY 6 VDOLKNYNDLVPELPAPEDVETNCENSAFSCFOKAOLKSANTGNERRIINVSIIKKLR 65
 Db 48 VSOVENKSTSKVP-----NNULPRCFPPHNLNRVFTQSGNOAISHSIKVLK 94
 QY 66 KPPST 70
 Db 95 HPCST 99

RESULT 14
 CCB3_MOUSE
 ID CCB3_MOUSE STANDARD; PRT; 484 AA.
 AC P54285;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE DIHYDROPYRIDINE-SENSITIVE L-TYPE, CALCIUM CHANNEL BETA-3 SUBUNIT
 DE (CAB3A/CAB3B) (CCNB3).
 GN CACNB3 OR CACNLB3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96039278; PubMed=7490102;
 RA Chin H., Kwon O.-J., Jung H.H., Kim D.S., Kozak C.A.;
 RT "Genetic mapping of the mouse genes encoding the voltage-sensitive
 RT calcium channel subunits";
 RL Genomics 28:592-595(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RX MEDLINE=96184890; PubMed=8617257;
 RA Murakami M., Wissenbach U., Flockerzi V.;
 RT "Gene structure of the murine calcium channel beta3 subunit, cDNA and
 RT characterization of alternative splicing and transcription
 RT products";
 RL Eur. J. Biochem. 236:138-143(1996).
 CC -!- FUNCTION: THIS PROTEIN IS A SUBUNIT OF THE DIHYDROPYRIDINE (DHP)
 CC SENSITIVE CALCIUM CHANNEL.
 CC -!- SUBUNIT: THE L-TYPE CALCIUM CHANNEL IS COMPOSED OF FOUR SUBUNITS:
 CC ALPHA-1, ALPHA-2, BETA AND GAMMA.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; CAB3A (SHOWN HERE) AND CAB3B;
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL BETA SUBUNITS FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U20372; AAA91105.1;
 DR EMBL; X94404; -; NOT_ANNOTATED_CDS.
 DR EMBL; X94406; -; NOT_ANNOTATED_CDS.
 DR EMBL; X94405; -; NOT_ANNOTATED_CDS.
 DR MGB; MGI:103307; Cacb3.
 DR InterPro; IPR000584; -;
 DR Pfam; PF00774; Ca_channel_B; 1.
 KW Ionic channel; Ion transport; Voltage-gated channel; Calcium channel;
 KW Glycoprotein; Phosphorylation; SH3 domain; Multigene family;
 KW Alternative splicing.
 FT DOMAIN 59 120
 FT CARBOHYD 67 67
 FT SEQUENCE 484 AA; 54562 MW; 730B1C2057230EB1 CRC64;
 SQ SEQUENCE 484 AA; 54562 MW; 730B1C2057230EB1 CRC64;

Query Match

11.18; Score 71.5; DB 1; Length 484;

Best Local Similarity 22.7%; Pred. No. 12;
 Matches 32; Conservative 19; Mismatches 43; Indels 47; Gaps 6;

QY 4 DIVDQLKNTVND-----LVPE-----FLPAPEDVETNCENSAFSCFOKAOLKSANTGNNE 53
 Db 91 DFLHIKEKYSNDWIGRLVKEGGAIAFIPSPORLE-----SIRLKQEQARRSGNPS 142
 QY 54 RIINVSIIKKLRKPPSTNAGRRKOKHR-----LTPCSCDSYEKKPPKEFL 97
 Db 143 SLGDIG---FRSPPPSLAKQKQKQAEHVPPYDVVPSMRPVLVGLPSLKGYE----- 191

QY 98 ERFKSLLOKMIHQHLSRTHG 118
 Db 192 --VTDMQKALDFLKHFDG 210

RESULT 15
 AHR_MOUSE
 ID AHR_MOUSE STANDARD; PRT; 805 AA.
 AC P30561;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE AH RECEPTOR PRECURSOR (ARYL HYDROCARBON RECEPTOR) (AHR).
 GN AHR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 10-25.
 RX MEDLINE=92231934; PubMed=1314586;
 RA Ema M., Sogawa K., Watanabe N., Chujoh Y., Matsushita N., Gotoh O.,
 RA Funae Y., Fujii-Kuriyama Y.;
 RT "cDNA cloning and structure of mouse putative Ah receptor.";
 RL Biochem. Biophys. Res. Commun. 184:246-253(1992).
 RN [2]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=C57BL;
 RX MEDLINE=92390411; PubMed=1325649;
 RA Burbach K.M., Poland A., Bradfield C.A.;
 RT "Cloning of the Ah-receptor cDNA reveals a distinctive
 RT ligand-activated transcription factor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:8185-8189(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DRA/2J; TISSUE=Liver;
 RX MEDLINE=95050471; PubMed=7961644;
 RA Ema M., Ohe N., Suzuki M., Mimura J., Sogawa K., Ikawa S.,
 RA Fujii-Kuriyama Y.;
 RT "Dioxin binding activities of polymorphic forms of mouse and human
 RT arylhydrocarbon receptors.";
 RL J. Biol. Chem. 269:27337-27343(1994).
 CC -!- FUNCTION: LIGAND-ACTIVATED TRANSCRIPTIONAL ACTIVATOR. BINDS TO THE
 CC ARE PROMOTER REGION OF GENES IT ACTIVATES. ACTIVATES THE
 CC EXPRESSION OF DRUG METABOLIZING ENZYMES GENES (SUCH AS THE CYP1A1
 CC GENE). IT MEDIATES BIOCHEMICAL AND TOXIC EFFECTS OF HALOGENATED
 CC AROMATIC HYDROCARBONS.
 CC -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
 CC BHLH PROTEIN. IN THE NUCLEUS, HETERODIMER OF AHR AND ARNT.
 CC -!- SUBCELLULAR LOCATION: FIRST CYTOPLASMIC, UPON BINDING WITH LIGAND
 CC AND INTERACTION WITH A HSP90, IT TRANSLOCATES TO THE NUCLEUS.
 CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS.
 CC -!- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

Result No.	Score	Query		Length	DB	ID	Description
		Match	No.				
1	99.5	15.4	114	4	Q9UBA3		Q9UBA3 homo sapien
2	99.5	15.4	136	4	Q00440		Q00440 homo sapien
3	91.5	14.1	567	11	Q05208		Q05208 mus musculu
4	89.5	13.8	162	6	Q9XS06		Q9XSJ6 ovis aries
5	83.5	12.9	3052	14	Q82933		Q82933 johnsongras
6	81.5	12.6	566	5	Q9VYE8		Q9VYE8 drosophila
7	75.5	11.7	336	11	Q62612		Q62612 rattus norv
8	75.5	11.7	566	11	Q62611		G62611 rattus norv
9	74	11.4	2197	5	Q96296		Q96296 plasmodium
10	73.5	11.4	187	13	Q9W756		Q9W756 gallus gall
11	73.5	11.4	868	5	Q9NAH8		Q9NAH8 caenorhabdi
12	73	11.3	848	4	Q75141		Q75141 homo sapien
13	72.5	11.2	305	11	Q9JJ48		Q9JJ48 mus musculu
14	72.5	11.2	1101	5	Q62471		Q62471 caenorhabdi
15	72	11.1	262	14	Q92434		Q92434 bombyx mori
16	72	11.1	424	5	Q23239		Q23239 calliphora
17	72	11.1	848	11	Q9QVY1		Q9QVY1 mus sp. aro
18	72	11.1	1175	4	Q13632		Q13632 homo sapien
19	72	11.1	1427	4	Q14207		Q14207 homo sapien

DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DE 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
GN INTERLEUKIN-15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Meazza R., Ferrini S.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y09908; CAA71044.1; -
SQ SEQUENCE 136 AA; 15259 MW; 494ED8DC5D364CD CRC64;

Query Match 15.4%; Score 99.5; DB 4; Length 136;
Best Local Similarity 25.9%; Pred. No. 0.0064;
Matches 30; Conservative 23; Mismatches 48; Indels 15; Gaps 5;

QY 3 IDIVDLKKNYNDLVPEF-----LPAPEDVETNCEWSAFSCF----QKAQLKSANTGNNE 53
DB 25 VNVISDLKK-IEDLTQSHMDIATLYTESDVHPSCKVTAMKCFLLLELQVISLESGDASIH 83
QY 54 RIINVSITKKLKKPPSTNAGRQRHRLTQPCSDSYKKPPKPEFLERFKSLQKMIH 109
DB 84 TVENLII--LANNSSNGNVTES---GCKECELEEKNIKEFLQSFVHIQVOMFIN 134

RESULT 3
Q05208 PRELIMINARY; PRT; 567 AA.
AC Q05208;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE ST2L PROTEIN PRECURSOR.
GN LY84.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=93170492; PubMed=7916701;
RX Yanagisawa K., Takagi T., Tsukamoto T., Tetsuka T., Tomimaga S.;
RT "Presence of a novel primary response gene ST2L, encoding a product
RL highly similar to the interleukin 1 receptor type 1.";
RL FEBS Lett. 318:83-87(1993).
DR EMBL: D13695; BAA02854.1; -.
DR MGD: MGI:98427; Ly84.
DR INTERPRO: IPR00157; -.
DR INTERPRO: IPR002052; -.
DR INTERPRO: IPR003006; -.
DR PFAM: PF00047; 19; 3.
DR PFAM: PF01582; TIR; 1.
DR PROSITE: PS00092; N6_MTASE; UNKNOWN_1.
KW SIGNAL.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 567 ST2L PROTEIN.
SQ SEQUENCE 567 AA; 64801 MW; 9228CE227D95B0BC CRC64;

Query Match 14.1%; Score 91.5; DB 11; Length 567;
Best Local Similarity 25.6%; Pred. No. 0.19;
Matches 33; Conservative 20; Mismatches 39; Indels 37; Gaps 7;

QY 7 DQKKNYNDLVPELPAPEDVETNCEWSAFSCFQKAQLKSANTGNNERIINVSIKLKKR 66
DB 76 DRLK-----FLPARVE-----DSGIYACV-----IRSPNL-NKTGLNVNTHK---K 113
QY 67 PPSTN-----AGRRQKRLTQPCSDSYKKPPKPEFLERFKSLQKMIHQH----- 111

DB 114 PPSCNIPDYLMYSTVRGSDKNFKITCTIDLYNWTAPVQVFNCKALOEPRFRAHRSYLF 173
QY 112 LSSRTHGE 120
DB 174 IDNVTHDE 182

RESULT 4
Q9XSJ6 PRELIMINARY; PRT; 162 AA.
AC Q9XSJ6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE INTERLEUKIN-15.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Casey G.J., Chaplin P.J.;
RT "Isolation of interleukin-15 mRNA transcripts from T and B cells
RT circulating in effluent lymph.";
RL J. Interferon Cytokine Res. 0:0-0(1999).
DR EMBL: AF149700; AAD37425.1; -.
SQ SEQUENCE 162 AA; 18510 MW; 87195650E45F1E5D CRC64;

Query Match 13.8%; Score 89.5; DB 6; Length 162;
Best Local Similarity 28.6%; Pred. No. 0.082;
Matches 24; Conservative 11; Mismatches 48; Indels 1; Gaps 1;

QY 26 DVETNCEWSAFSCFQKAQLKSANTGNNERIINVSIKLKKRPPSTNAGRQKHLTCPCSC 85
DB 78 DAHPNCKVTAQCFLLELRVLHESKNAAIYEI-IENLTMLADRNLSIENKTELGCCKE 136
QY 86 DSYERKPPKPEFLERFKSLQKMIH 109
DB 137 EELEKKSKEFLKSFVHIQVOMFIN 160

RESULT 5
Q82933 PRELIMINARY; PRT; 3052 AA.
AC Q82933;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE JOHNSON GRASS MOSAIC VIRUS PROTEASE 1 AND 3, HELPER COMPONENT 6K
DE PROTEIN, COAT PROTEIN, NUCLEAR INCLUSION PROTEINS.
OS Johnsongrass mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=31742;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94200956; PubMed=81505099;
RA Gough K.H., Shukla D.D.;
RT "Nucleotide sequence of Johnsongrass mosaic potyvirus genomic RNA.";
RL Intervirology 36:181-192(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Gough K.H.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Nurhayati E.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z26920; CAA81549.1; -.
DR MEROPS; C04.001; -.
DR MEROPS; C06.001; -.

```

DR MEROPS: S30.001: -
DR INTERPRO: IPR001205: -
DR INTERPRO: IPR001410: -
DR INTERPRO: IPR001456: -
DR INTERPRO: IPR001592: -
DR INTERPRO: IPR001730: -
DR INTERPRO: IPR002540: -
DR PFAM: PF00270: DEAD; 1.
DR PFAM: PF00680: RNA_dep_RNA_pol; 1.
DR PFAM: PF00767: Poty_coat; 1.
DR PFAM: PF00851: Peptidase_C6; 1.
DR PFAM: PF00863: Peptidase_C4; 1.
DR PFAM: PF01577: Poly_P1; 1.
DR PRINTS: PR00966: NIAPOTYPTASE.
KW Coat protein; Protease.
FT CHAIN 239 699
FT FT CHAIN 700 1046
FT FT CHAIN 1047 1098
FT FT CHAIN 1099 1757
FT FT CHAIN 1758 1803
FT FT CHAIN 1804 1991
FT FT CHAIN 1992 2233
FT FT CHAIN 2234 2749
FT FT CHAIN 2750 3052
FT FT CHAIN 3052 347249 MW; 08CD8831A73EBCA9 CRC64;
SQ SEQUENCE 3052 AA; 347249 MW; 08CD8831A73EBCA9 CRC64;

Query Match 12.9%; Score 83.5; DB 14; Length 3052;
Best Local Similarity 23.1%; Pred. No. 7.7; Indels 3; Gaps 3;
Matches 25; Conservative 23; Mismatches 57;

QY 5 IVDOLKYNVNDLPEFLPAPEDVETNCWFSQFQAQLKSANTGNININVSIKLKL 64
DB 220 LVNALQYDEED-VKQICHYSFDEAERAFWKFTENHTAQRREADHTNHPV-MSVEECG 277
QY 65 RKPSPNAGRQRHRTCTPCSC-DSYEKKPKPEFLERFKSLQKMIHQH 111
DB 278 RRAAMLENAPHOGFKITCKHCFQTFDEHSDSEVCERLHNALRIEEQN 325

RESULT 6
QYVE8 PRELIMINARY; PRT; 566 AA.
AC Q9VVE8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE CG15745 PROTEIN.
GN CG15745.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

```

```

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003492; AAF48250.1; -.
DR FLYBASE: FBgn0030469; CG15745.
SQ SEQUENCE 566 AA; 60620 MW; 31FA41FA4AEF427F CRC64;

```

```

Query Match 12.6%; Score 81.5; DB 5; Length 566;
Best Local Similarity 26.1%; Pred. No. 2.1;
Matches 24; Conservative 21; Mismatches 40; Indels 7; Gaps 3;

```

```

QY 26 DVETNCWSAFS-CFQKAQLKSANTGNININVSIIKLRKPPPTNAGRRQKRLTCP 83
DB 120 DIGYPCVEVASISELAKRAQLKRAQFFGNVG----GLARDSETSTTRITRTNRSAYP 175
QY 84 SCDSYEKKPKPEFLERFKSL-QKMIHQHLS 114
DB 176 SCKTERGPKVQQLIDQFQAMIVQQQQQQLSN 207

```

```

RESULT 7
Q62612 PRELIMINARY; PRT; 336 AA.
AC Q62612;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE FIT-1S PRECURSOR.
GN FIT-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94178260; PubMed=8131748;
RA Bergers G., Reikerstorfer A., Bräselmann S., Graninger P.,
RA Busslinger M.;
RT "Alternative promoter usage of the Fos-responsive gene Fit-1 generates
RT mRNA isoforms coding for either secreted or membrane-bound proteins
RT related to the IL-1 receptor.";
RL EMO J. 13:1176-1188(1994).
DR EMBL: U04319; AAA67172.1; -.
DR INTERPRO: IPR003006; -.
PFAM: PF00047; Ig; 3.
KW SIGNAL.
FT SIGNAL 1 26
FT CHAIN 27 336
SQ SEQUENCE 336 AA; 38090 MW; 61C6A6FA8FE319D2 CRC64;

```

```

Query Match      11.7%; Score 75.5; DB 11; Length 336;
Best Local Similarity 22.2%; Pred. No. 5;
Matches 30; Conservative 26; Mismatches 42; Indels 37; Gaps 6;

QY 3 IDVDOLKNNYNDLVP-----EFLPAPEDVETNCWSAFSCFOKAOLKSAN 48
DB 49 INPVEWYYSNTNERIPTOKRNRIFVSRDLKFLPAKVE-----DSGIYTCVIRSP-ESIK 102

QY 49 TGNNERIINVSIIKKLRKPP-----STNAGRRQKHRLTCPCSDSYEKKPKPEFLE 98
DB 103 TGS-----LNVYIYK---RPPNCKIPDYMSTVDGSDKNSKITCPTIALYNNWTPVQWPK 155

QY 99 RFKSLQRMHIOHLS 113
DB 156 NCKALQGPFRFRAHMS 170

RESULT 8
Q62611 PRELIMINARY; PRT; 566 AA.
AC Q62611;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE FIT-1M PRECURSOR.
GN FIT-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94178260; PubMed=8131748;
RA Bergers G., Reikerstorfer A., Braselmann S., Graninger P.,
RA Busslinger M.;
RT "Alternative promoter usage of the Fos-responsive gene Fit-1 generates
RT mRNA isoforms coding for either secreted or membrane-bound proteins
RT related to the IL-1 receptor.";
RL EMO J. 13:1176-1188(1994).
DR EMBL; U04317; AAA18480.1; -.
DR INTERPRO; IPR00157; -.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; 19; 3.
DR PFAM; PF01582; TIR; 1.
KW Signal.
FT SIGNAL.
FT CHAIN 27 566
FT CHAIN 27 566
SQ SEQUENCE 566 AA; 64405 MW; 50554A29437A0C31 CRC64;

Query Match      11.7%; Score 75.5; DB 11; Length 566;
Best Local Similarity 22.2%; Pred. No. 8.6;
Matches 30; Conservative 26; Mismatches 42; Indels 37; Gaps 6;

QY 3 IDVDOLKNNYNDLVP-----EFLPAPEDVETNCWSAFSCFOKAOLKSAN 48
DB 49 INPVEWYYSNTNERIPTOKRNRIFVSRDLKFLPAKVE-----DSGIYTCVIRSP-ESIK 102

QY 49 TGNNERIINVSIIKKLRKPP-----STNAGRRQKHRLTCPCSDSYEKKPKPEFLE 98
DB 103 TGS-----LNVYIYK---RPPNCKIPDYMSTVDGSDKNSKITCPTIALYNNWTPVQWPK 155

QY 99 RFKSLQRMHIOHLS 113
DB 156 NCKALQGPFRFRAHMS 170

RESULT 9
O96296 PRELIMINARY; PRT; 2197 AA.
AC O96296;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

```

```

DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE PFEMP1.
GN PFBI055C.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99021743; PubMed=9804551;
RA Gardner M.J., Tettelin H., Carucci D.J., Cummings L.M., Aravind L.,
RA Koonin E.V., Shalloom S., Mason T., Yu K., Fujii C., Pederson J.,
RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M.,
RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Science 282:1126-1132(1998).
DR EMBL; AE001434; AAC71996.1; -.
SQ SEQUENCE 2197 AA; 249668 MW; D2A92DB54535C143 CRC64;

Query Match      11.4%; Score 74; DB 5; Length 2197;
Best Local Similarity 29.5%; Pred. No. 52;
Matches 28; Conservative 16; Mismatches 41; Indels 10; Gaps 4;

QY 9 LKNYVNDLVPFLPAPEDVETNCWSAF--SCFOKAOLKSANTGNRIINVSIIKKLRK 66
DB 1594 VKSFLETWIPK-TAVVNDQDNVTKLSFGSGCSAISTN-GNEEADICMIKKLEKK 1651
QY 67 -----PPSTNAGRRQKHRLTCPCSDSYEKKPKPE 95
DB 1652 IDECKRKPGENSGOTCNETLTHPLDVQDEPLEE 1686

RESULT 10
Q9W756 PRELIMINARY; PRT; 187 AA.
AC Q9W756;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (INTERLEUKIN 15 PRECURSOR).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCWL; TISSUE=LIVER;
RC Burnside J., Sofer L.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SC; TISSUE=LIVER;
RC Choi K.D., Lillehoj H.S., Burnside J.;
RT "Gallus gallus mRNA for IL-15 precursor.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152927; AAD38392.1; -.
DR EMBL; AF139097; AAF61446.1; -.
SQ SEQUENCE 187 AA; 21964 MW; 470601EBF8837095 CRC64;

Query Match      11.4%; Score 73.5; DB 13; Length 187;
Best Local Similarity 25.0%; Pred. No. 4.3;
Matches 24; Conservative 11; Mismatches 34; Indels 27; Gaps 4;

QY 25 EDVETNCWSAFSCF---OKAOLKSANTGNRIINVSIIKKLRKPPSTNAGRRQKHRLT 81
DB 98 EDIE--CQEPVMRCFFLEMKVILHECD-----IKCSRKHVNRNWKGNARFA 144
QY 82 -----CPSCDSYEKKPKPEFLERFKSLQK 106

```

Db 145 TYQLNSTAKKCKCEYEYKNTFTFIQSPVKVIQR 180

RESULT 11

Q9NAH8 PRELIMINARY; PRT; 868 AA.

ID Q9NAH8

AC Q9NAH8

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)

DE Y47D3A.14 PROTEIN.

GN Y47D3A.14

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RA SEQUENCE FROM N.A.

RA Matthews L.;

RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RN SEQUENCE FROM N.A.

RA none;

RT "Genome sequence of the nematode C.elegans: A platform for

RT investigating biology."

RL Science 282:2012-2018(1998).

DR EMBL; AL117202; CAB55073.1; -.

SQ SEQUENCE 868 AA; 100165 MW; 09A2A97066E373D7 CRC64;

Query Match 11.4%; Score 73.5; DB 5; Length 868;

Best Local Similarity 29.2%; Pred. No. 22;

Matches 19; Conservative 11; Mismatches 30; Indels 5; Gaps 1;

QY 33 WSAFSCFQKALQS-----ANTGNNEIRIIVNSIKLKRKPPSTNAGRQKRLTCPSCDS 87

Db 680 WSRVQAFRRKNNKSGGGTNSDSSESEKVLKLRARRPSDETIVPVPYHIVCLFPS 739

QY 88 YKKP 92

Db 740 HVAIP 744

RESULT 12

O75141 PRELIMINARY; PRT; 848 AA.

ID O75141

AC O75141

DT 01-NOV-1998 (TREMBlrel. 08, Created)

DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)

DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)

DE KIAA0650 PROTEIN (FRAGMENT).

GN KIAA0650.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RA SEQUENCE FROM N.A.

RC TISSUE=BRAIN;

RX MEDLINE=98403880; PubMed=9734811;

RX Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,

RA Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. X.

RT The complete sequences of 100 new cDNA clones from brain which can

RT code for large proteins in vitro."

RL DNA Res. 5:169-176(1998).

DR EMBL; AB014550; BAA31625.1; -.

FT NON_TER 1

SQ SEQUENCE 848 AA; 95756 MW; 2192522F40ED9E7B CRC64;

Query Match 11.3%; Score 73; DB 4; Length 848;

Best Local Similarity 25.5%; Pred. No. 24;

Matches 27; Conservative 22; Mismatches 47; Indels 10; Gaps 5;

QY 19 EFLPAPEDVETNC-EWSAFSCFQKALQSANTGNNEIRIIVNSIKLKRKPPSTNAGRQK 77

Db 66 KPIPGPGNKDLCTWREFSDFIRVLIS--GPPAKLLLDWPELKESIPVIN-GRDLQ 121

QY 78 HRLTCPSCDSYEKKPKPEFLERFKSLQKMIHOHL--SSRTHGSED 121

Db 122 NPIIVQLCDOWNPAP---VOHVKISLTKASNLKLMPSNOQHKTE 164

RESULT 13

Q9JU48 PRELIMINARY; PRT; 305 AA.

ID Q9JU48

AC Q9JU48

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)

DE PUTATIVE ZINC FINGER PROTEIN FLIZ1.

GN FLIZ1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RN SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=LIVER;

RA Dahn K.E., Mueller A.M.;

RT "Fetal liver zinc finger 1 (Fliz1), a new gene transcribed in the

RT hematopoietic stem and progenitor cell fraction of the E12 murine

RT fetal liver containing three repeats of a C3H zinc finger motif."

RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF061961; AAF74513.1; -.

SQ SEQUENCE 305 AA; 34922 MW; 2F45A0758D56245C CRC64;

Query Match 11.2%; Score 72.5; DB 11; Length 305;

Best Local Similarity 26.3%; Pred. No. 9.1;

Matches 31; Conservative 14; Mismatches 44; Indels 29; Gaps 5;

QY 8 QLKNYVNDLVPELPAPEDVETNCWSAFSCFQKALQSANTGNNEIRIIVNSIKLKRKP 67

Db 107 QAKEMANAQPSLLPEPVKAGAGCTQAKQKN--KSKAGHKK---VKOKMKRW 160

QY 68 PST-----NAGRRQKRLTCPSCDSYEKKPKPEFLERFKSLQKMIHOHLSSR 115

Db 161 PETGDKGRALLKNSGREQ-----TDEPEKQPR-----VRMSQGFINGHTVER 205

RESULT 14

O62471 PRELIMINARY; PRT; 1101 AA.

ID O62471

AC O62471

DT 01-AUG-1998 (TREMBlrel. 07, Created)

DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)

DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)

DE Y45F10B.10 PROTEIN.

GN Y45F10B.10.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RN SEQUENCE FROM N.A.

RA Mcmurray A.;

RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RN SEQUENCE FROM N.A.

RX MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,

RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,

RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,

RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,

```

RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
DR EMBL; AL021487; CAAL16357.1; -.
DR INTERPRO; IPR001680; -.
DR PFAM; PF00400; WD40; 9.
DR PRINTS; PR00320; GPROTEINBRPT.
DR PROSITE; PS00678; WD_REPEATS; UNKNOWN_2.
SQ SEQUENCE 1101 AA; 121014 MW; 9C9C657F8A67D6B0 CRC64;

Query Match 11.2%; Score 72.5; DB 5; Length 1101;
Best Local Similarity 31.5%; Pred. No. 36;
Matches 23; Conservative 14; Mismatches 23; Indels 13; Gaps 4;

QY 40 OKAQLKSANTGNRIINVSIIKKLRKP--PSTNAGRQK-HRLTCPCSDSYEKPPKEF 96
DB 1013 RSARQSVSSASNEPVASTSAGEIKKDPILSSNNGNAQSAPRATAP-----KPTFDM 1065

QY 97 LERFK---SLLOK 106
DB 1066 LERSKRTSLIEK 1078

RESULT 15
O92434
ID O92434 PRELIMINARY; PRT; 262 AA.
AC O92434;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE ACNVPV ORF69.
OS Bombyx mori nuclear polyhedrosis virus (BmNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=10458;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T3;
RX MEDLINE=97329351; PubMed=9185864;
RA Kamita S.G., Maeda S.;
RT "Sequencing of the putative DNA helicase-encoding gene of the Bombyx
RT mori nuclear polyhedrosis virus and fine-mapping of a region involved
RT in host range expansion.";
RL Gene 190:173-179(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=T3;
RA Gomi S., Majima K., Maeda S.;
RT "Sequence analysis of the genome of Bombyx mori
RT nucleopolyhedrovirus.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; L31180; AAC63742.1; -.
DR INTERPRO; IPR002877; -.
DR PFAM; PF01728; FtsJ; 1.
SQ SEQUENCE 262 AA; 30384 MW; E3AF85E1B02DA48C CRC64;

Query Match 11.1%; Score 72; DB 14; Length 262;
Best Local Similarity 21.7%; Pred. No. 8.7;
Matches 18; Conservative 19; Mismatches 36; Indels 10; Gaps 2;

QY 30 NCEWSAFSCFOKAQLKSANTGNRIINVSIIKKLRKPSTNAGRQKH-----RLTCPS 84
DB 176 NCVLKVDFAEHKTQMLN-----KFNHFKEKVLKPPSSRPANSERYLICENKLVRY 230

QY 85 CDSYEKKPPKEFLERFKSLLOQM 107

```

```

Db 231 CNDYVNELEKQFKKYRIQLKNL 253

Search completed: May 23, 2001, 11:19:43
Job time: 596 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:11:34 ; Search time 108.07 Seconds
(without alignments)
8.463 Million cell updates/sec

Title: US-09-522-217-2_COPY_41_56
Perfect score: 80
Sequence: 1 QLIDVDQLKNVNDL 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	100.0	40	B18625	Antigenic peptide
2	80	100.0	162	B18623	A human zalphall 1
3	80	100.0	519	B18627	Amino acid sequenc
4	56	70.0	146	B18624	A mouse zalphall 1
5	56	70.0	510	B18628	Amino acid sequenc
6	41	51.2	812	20 Y35560	C. pneumoniae prot
7	40.5	50.6	600	20 Y29861	Human secreted pro
8	40.5	50.6	704	20 Y29517	Human lung tumour
9	40.5	50.6	704	21 B44467	Human lung tumour
10	40	50.0	171	20 Y35668	Chlamydia pneumoni
11	40	50.0	478	21 Y95040	Candida albicans p

12	40	50.0	541	19 W81514	Chromodomain y b (
13	40	50.0	554	19 W81513	Chromodomain y a (
14	40	50.0	934	17 R92289	Cycloisomaltoligo
15	40	50.0	972	17 R92288	Cycloisomaltoligo
16	40	50.0	1091	20 Y30337	Bacillus CTase pr
17	40	50.0	1091	20 Y30337	protein encoded by
18	39	48.8	449	19 W98600	H. pylori GHPO 128
19	39	48.8	805	16 R80561	Murine Ah receptor
20	38	47.5	102	21 B09952	Rice disease resis
21	38	47.5	133	20 Y40120	Amino acid sequenc
22	38	47.5	245	20 Y20027	B. burgdorferi ant
23	38	47.5	259	22 B63752	Human prostate can
24	38	47.5	264	19 W42388	Aquifex VF-5 alkali
25	38	47.5	265	20 Y20026	B. burgdorferi ant
26	38	47.5	389	21 Y85619	Thermotable amino
27	38	47.5	389	21 Y77535	Pyrococcus thermos
28	38	47.5	440	19 W26782	Mouse Friend virus
29	38	47.5	459	19 W26781	Mouse Friend virus
30	38	47.5	572	19 W69391	Aspergillus fumiga
31	38	47.5	716	19 W69846	Amino acid sequenc
32	38	47.5	918	20 Y35654	Amino acid sequenc
33	38	47.5	1247	21 B18215	Plasmodium falcipa
34	37.5	46.9	434	21 B41968	Human ORFX ORF1732
35	37	46.2	28	21 Y88726	Core polypeptide f
36	37	46.2	52	20 W89859	Antigen 2 from ciu
37	37	46.2	65	20 Y05837	Banana fruit-assoc
38	37	46.2	163	21 G12749	Arabidopsis thalia
39	37	46.2	178	21 G21975	Arabidopsis thalia
40	37	46.2	203	21 G21974	Arabidopsis thalia
41	37	46.2	204	20 Y34586	C. pneumoniae prot
42	37	46.2	283	20 Y26895	Vibrio cholerae th
43	37	46.2	283	21 Y59126	Amino acid sequenc
44	37	46.2	359	21 G41243	Arabidopsis thalia
45	37	46.2	417	21 G41242	Arabidopsis thalia

ALIGNMENTS

RESULT 1
B18625
ID B18625 standard; Peptide; 40 AA.
XX
AC B18625;
XX
DT 22-JAN-2001 (first entry)
XX
DE Antigenic peptide derived from a human zalphall ligand polypeptide.
XX
KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
XX
OS Homo sapiens.
XX
FN WO200053761-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-US06067.
XX
PR 09-MAR-1999; 99US-0264908.
PR 11-MAR-1999; 99US-0265992.
PR 01-JUL-1999; 99US-0142013.
XX (ZYMO) ZYMOGENETICS INC.
XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX WPI; 2000-565600/52.
XX New human cytokine, designated zalphall ligand, useful for stimulating
PT the proliferation and/or development of haematopoietic cells in vitro

PT and in vivo, and for treating tumourigenesis -
 XX Example 34; Page 227; 256pp; English.
 XX

CC The present sequence was used to raise antibodies, and is derived from
 CC a human zalphall ligand polypeptide. zalphall ligand is a cytokine.
 CC The zalphall ligand is useful for stimulating the proliferation and
 CC development of haematopoietic cells in vitro and in vivo. zalphall
 CC ligand polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 XX Sequence 40 AA;
 SQ

Query Match 100.0%; Score 80; DB 21; Length 40;
 Best Local Similarity 100.0%; Pred. No. 8.7e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLIDIVDQLKNYVNDL 16
 DB 10 qlidivdqlknyvndl 25
 |||||

RESULT 2
 B18623
 ID B18623 standard; Protein; 162 AA.
 XX
 AC B18623;
 XX
 DT 22-JAN-2001 (first entry)
 XX
 DE A human zalphall ligand polypeptide.
 XX
 KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX
 OS Homo sapiens.
 XX
 PN WO200053761-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06067.
 XX
 PR 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 XX
 DR WPI; 2000-565600/52.
 DR N-PSDB; A75552.
 XX
 XX New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX
 PS Disclosure; Page 205-206; 256pp; English.
 XX

CC The present sequence represents a human zalphall ligand polypeptide,
 CC which is a cytokine. The zalphall ligand is useful for stimulating the
 CC proliferation and development of haematopoietic cells in vitro and in
 CC vivo. zalphall ligand polynucleotides can be used as primers or probes
 CC for cloning the zalphall gene. The zalphall ligand is useful for
 CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
 CC used for treating leukaemias and lymphomas. Antagonists against zalphall
 CC ligand are useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 XX Sequence 162 AA;
 SQ

Query Match 100.0%; Score 80; DB 21; Length 162;
 Best Local Similarity 100.0%; Pred. No. 4.1e-05;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps -0;

QY 1 QLIDIVDQLKNYVNDL 16
 DB 41 qlidivdqlknyvndl 56
 |||||

RESULT 3
 B18627
 ID B18627 standard; Protein; 519 AA.
 XX
 AC B18627;
 XX
 DT 22-JAN-2001 (first entry)
 XX
 DE Amino acid sequence of MBP-human zalphall ligand fusion in pTAP126.
 XX
 KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO200053761-A2.
 XX
 PD 14-SEP-2000.
 XX
 PF 09-MAR-2000; 2000WO-US06067.
 XX
 PR 09-MAR-1999; 99US-0264908.
 PR 11-MAR-1999; 99US-0265992.
 PR 01-JUL-1999; 99US-0142013.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
 XX
 DR WPI; 2000-565600/52.
 DR N-PSDB; A75599.
 XX
 XX New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX
 PS Example 31; Page 233-235; 256pp; English.
 XX
 CC The present sequence represents a MBP-human zalphall ligand fusion in
 CC the plasmid pTAP126. zalphall ligand is a cytokine. The zalphall ligand
 CC is useful for stimulating the proliferation and development of
 CC haematopoietic cells in vitro and in vivo. Zalphall ligand

CC polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.

XX Sequence 519 AA;

Query Match 100.0%; Score 80; DB 21; Length 519;
 Best Local Similarity 100.0%; Pred. No. 0.00015;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QLIDIVDQLKNYVNDL 16
 |||||
 Db 398 qlidivdqlknyvndl 413

RESULT 4

ID B18624 B18624 standard; Protein; 146 AA.

AC B18624;

XX 22-JAN-2001 (first entry)

XX A mouse zalphall ligand polypeptide.

XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.

XX Mus musculus.

XX WO200053761-A2.

XX 14-SEP-2000.

XX 09-MAR-2000; 2000WO-US06067.

XX 09-MAR-1999; 99US-0264908.

XX 11-MAR-1999; 99US-0265992.

XX 01-JUL-1999; 99US-0142013.

XX (ZYMO) ZYMOGENETICS INC.

XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;

XX WPI; 2000-565600/52.

XX N-PSDB; A75580.

XX New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -

XX Disclosure; Page 222-223; 256pp; English.

XX The present sequence represents a mouse zalphall ligand polypeptide,
 CC which is a cytokine. The zalphall ligand is useful for stimulating the
 CC proliferation and development of haematopoietic cells in vitro and in
 CC vivo. Zalphall ligand polynucleotides can be used as primers or probes
 CC for cloning the zalphall gene. The zalphall ligand is useful for
 CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
 CC used for treating leukaemias and lymphomas. Antagonists against zalphall
 CC ligand are useful as research reagents for characterizing ligand-receptor

CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.

XX Sequence 146 AA;

Query Match 70.0%; Score 56; DB 21; Length 146;
 Best Local Similarity 80.0%; Pred. No. 0.18;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LIDIVDQLKNYVNDL 16

|||||

Db 35 lidiveqlklyendl 49

RESULT 5

ID B18628 B18628 standard; Protein; 510 AA.

XX B18628;

XX 22-JAN-2001 (first entry)

XX Amino acid sequence of MBP-mouse zalphall ligand fusion in pTAP134.

XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.

XX Synthetic.

XX Mus musculus.

XX WO200053761-A2.

XX 14-SEP-2000.

XX 09-MAR-2000; 2000WO-US06067.

XX 09-MAR-1999; 99US-0264908.

XX 11-MAR-1999; 99US-0265992.

XX 01-JUL-1999; 99US-0142013.

XX (ZYMO) ZYMOGENETICS INC.

XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;

XX WPI; 2000-565600/52.

XX N-PSDB; A75602.

XX New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -

XX Example 31; Page 239-240; 256pp; English.

XX The present sequence represents a MBP-mouse zalphall ligand fusion in
 CC the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
 CC is useful for stimulating the proliferation and development of
 CC haematopoietic cells in vitro and in vivo. Zalphall ligand
 CC polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to

CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 SQ Sequence 510 AA;

Query Match 70.0%; Score 56; DB 21; Length 510;
 Best Local Similarity 80.0%; Pred. No. 0.7;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LIDIVDQLKNVNDL 16
 |||||:|||||
 Db 399 lidiveqikiyendl 413

RESULT 6
 Y35560
 ID Y35560 standard; Protein; 812 AA.
 AC Y35560;

DT 13-SEP-1999 (first entry)

XX C. pneumoniae protein involved in cell wall synthesis.

XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
 KW vaccine; neutralising epitope.

XX Chlamydia pneumoniae.

XX WO9927105-A2.

XX 03-JUN-1999.

XX 20-NOV-1998; 98WO-IB01890.

XX 04-NOV-1998; 98US-0107078.

PR 21-NOV-1997; 97FR-0014673.

XX (GEST) GENSET.

PA Griffais R;

XX WPI: 1999-357842/30.

XX Genome sequence of Chlamydia pneumoniae

PS Page 1304-1306; Disclosure; 1912pp; English.

XX Y34584-Y35879 represent the proteins encoded by all the open reading
 CC frames in the complete genome (see X91990) of Chlamydia pneumoniae.
 CC C. pneumoniae causes respiratory disease such as pneumonia and
 CC bronchitis and is thought to be a contributing factor in heart
 CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
 CC nodosum or pharyngitis. The polypeptides encoded by the open reading
 CC frames of the C. pneumoniae genome (see Y34584-Y35879) can be used in
 CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
 CC nucleotide sequences can also be used as immunogenic compositions,
 CC especially where the vector directs the expression of a neutralising
 CC epitope of C. pneumoniae.

XX Sequence 812 AA;

Query Match 51.2%; Score 41; DB 20; Length 812;
 Best Local Similarity 70.0%; Pred. No. 2.4e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 DIVDQLKNV 13

Db 417 divdylrnyl 426
 |||||:|:|:

RESULT 7
 Y29861
 ID Y29861 standard; Protein; 600 AA.

XX Y29861;

XX 17-NOV-1999 (first entry)

XX Human secreted protein clone cb98_4.

XX Human; secreted protein; biological activity; nutritional; cytokine;
 KW cell proliferation; differentiation; immune stimulating; vaccine;
 KW haematopoiesis regulation; tissue growth; haemostatic; thrombolytic;
 KW anti-inflammatory; tumour inhibition.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH Misc-difference 99

ET /note= "unspecified"

XX WO9946287-A1.

XX 16-SEP-1999.

XX 11-MAR-1999; 99WO-US05243.

XX 11-MAR-1998; 98US-0077521.

PR 14-MAY-1998; 98US-0079124.

PR 10-MAY-1999; 99US-0266105.

XX (GEMY) GENETICS INST INC.

XX Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Agostino MJ, Steininger RJ;

XX WPI: 1999-551362/46.

XX N-PSDB; Z21093.

XX Polynucleotides encoding secreted human proteins, derived from human
 PT fetal brain, human adult blood, human adult bladder, or human adult
 PT neural tissue cDNA libraries.

XX Claim 9; Page 99-101; 118pp; English.

XX Z21093 to Z21102 encode new human secreted proteins and Y29861 to Y29873
 CC represent the secreted proteins encoded by the polynucleotide sequences.
 CC Z21103 to Z21112 represent probes for the secreted proteins. The
 CC polynucleotides and proteins are predicted to have biological activities
 CC which would make them suitable for treating, preventing or ameliorating
 CC medical conditions in humans and animals, although no supporting data
 CC is given. Suggested activities include nutritional activity, cytokine
 CC and cell proliferation/differentiation activity, immune stimulating
 CC (e.g. as vaccines) or suppressing activity, haematopoiesis regulating
 CC activity, tissue growth activity, activin/inhibin activity,
 CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
 CC activity, receptor/ligand activity, anti-inflammatory activity,
 CC cadherin/tumour invasion suppressor activity, and tumour inhibition
 CC activity. The polynucleotides and proteins can also be used as
 CC nutritional sources or supplements. Such uses include use as a protein or
 CC amino acid supplement, use as a carbon source, use as a nitrogen source
 CC and use as a source of carbohydrate. They may also have utility in
 CC compositions used for bone, cartilage, tendon, ligament, and/or nerve
 CC tissue growth or regeneration, as well as for wound healing and tissue
 CC repair and replacement, and in the treatment of burns, incisions and
 CC ulcers. The proteins which induce cartilage and/or bone growth in
 CC circumstances where bone is not normally formed, have application in
 CC the healing of bone fractures and cartilage damage or defects in humans
 CC and other animals.

```

XX SQ Sequence 600 AA;
Query Match 50.6%; Score 40.5; DB 20; Length 600;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

QY 1 QLIDIVDQLKNYVNDL 16
::|::|::|::|::|
Db 192 ritdvlqg-knyveel 206

RESULT 8
Y29517
ID Y29517 standard; Protein; 704 AA.
XX AC Y29517;
XX DT 13-OCT-1999 (first entry)
XX DE Human lung tumour protein SAL-82 predicted amino acid sequence.
XX KW Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine;
XX KW immunotherapy; detection; inhibition.
XX OS Homo sapiens.
XX PN WO9938973-A2.
XX PD 05-AUG-1999.
XX PF 26-JAN-1999; 99WO-US01642.
XX PR 22-DEC-1998; 98US-0219245.
XX PR 28-JAN-1998; 98US-0015022.
XX PR 28-JAN-1998; 98US-0015029.
XX PR 18-MAR-1998; 98US-0040828.
XX PR 18-MAR-1998; 98US-0040831.
XX PR 23-JUL-1998; 98US-0122191.
XX PR 23-JUL-1998; 98US-0122192.
XX PA (CORI-) CORIXA CORP.
XX PI Fridakis TN, Lodes MJ, Mohamath R, Reed SG;
XX WPI: 1999-479187/40.
XX DR N-PSDB; Z07196.
XX PT Lung tumour specific polynucleotides for inhibiting the development
XX of lung cancer
XX PS Claim 3; Page 155-156; 171pp; English.
XX CC The present invention describes lung tumour specific polynucleotides
XX and tumour antigens. Z07144 to Z07246 and Z08301 to Z08325 represent
XX specifically claimed polynucleotides, and Y29486 to Y29571 represent
XX amino acid sequences from the present invention. The lung tumour
XX specific polynucleotides and polypeptides can be used in pharmaceutical
XX compositions and vaccines to inhibit the development of lung cancer.
XX They can also be used to detect lung cancer in a patient. Probes and
XX antibodies derived from the lung tumour sequences are useful in
XX detection of lung cancer..
XX SQ Sequence 704 AA;

Query Match 50.6%; Score 40.5; DB 20; Length 704;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

QY 1 QLIDIVDQLKNYVNDL 16
::|::|::|::|::|
Db 192 ritdvlqg-knyveel 206

RESULT 8
Y29517
ID Y29517 standard; Protein; 704 AA.
XX AC Y29517;
XX DT 13-OCT-1999 (first entry)
XX DE Human lung tumour protein SAL-82 predicted amino acid sequence.
XX KW Human; lung tumour protein; therapy; diagnosis; lung cancer; vaccine;
XX KW immunotherapy; detection; inhibition.
XX OS Homo sapiens.
XX PN WO9938973-A2.
XX PD 05-AUG-1999.
XX PF 26-JAN-1999; 99WO-US01642.
XX PR 22-DEC-1998; 98US-0219245.
XX PR 28-JAN-1998; 98US-0015022.
XX PR 28-JAN-1998; 98US-0015029.
XX PR 18-MAR-1998; 98US-0040828.
XX PR 18-MAR-1998; 98US-0040831.
XX PR 23-JUL-1998; 98US-0122191.
XX PR 23-JUL-1998; 98US-0122192.
XX PA (CORI-) CORIXA CORP.
XX PI Fridakis TN, Lodes MJ, Mohamath R, Reed SG;
XX WPI: 1999-479187/40.
XX DR N-PSDB; Z07196.
XX PT Lung tumour specific polynucleotides for inhibiting the development
XX of lung cancer
XX PS Claim 3; Page 155-156; 171pp; English.
XX CC The present invention describes lung tumour specific polynucleotides
XX and tumour antigens. Z07144 to Z07246 and Z08301 to Z08325 represent
XX specifically claimed polynucleotides, and Y29486 to Y29571 represent
XX amino acid sequences from the present invention. The lung tumour
XX specific polynucleotides and polypeptides can be used in pharmaceutical
XX compositions and vaccines to inhibit the development of lung cancer.
XX They can also be used to detect lung cancer in a patient. Probes and
XX antibodies derived from the lung tumour sequences are useful in
XX detection of lung cancer..
XX SQ Sequence 704 AA;

Query Match 50.6%; Score 40.5; DB 21; Length 704;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

QY 1 QLIDIVDQLKNYVNDL 16
::|::|::|::|::|
Db 192 ritdvlqg-knyveel 310

RESULT 10
Y35668
ID Y35668 standard; Protein; 171 AA.
XX AC Y35668;
XX DT 13-SEP-1999 (first entry)
XX DE Chlamydia pneumoniae transport polypeptide.
XX KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
XX sinusitis; purulent otitis media; erythema nodosum; pharyngitis;

```

```

Db 296 ritdvlqg-knyveel 310

RESULT 9
B44467
ID B44467 standard; Protein; 704 AA.
XX AC B44467;
XX DT 05-FEB-2001 (first entry)
XX DE Human lung tumour-specific antigen encoded by CDNA #103.
XX KW Lung tumour protein; lung cancer; cytostatic; vaccine.
XX OS Homo sapiens.
XX PN WO200060077-A2.
XX PD 12-OCT-2000.
XX PF 30-MAR-2000; 2000WO-US08560.
XX PR 02-APR-1999; 99US-0285323.
XX PR 09-AUG-1999; 99US-0370838.
XX PR 30-DEC-1999; 99US-0476235.
XX PR 03-MAR-2000; 2000US-0518809.
XX PA (CORI-) CORIXA CORP.
XX PI Reed SG, Lodes MJ, Mohamath R, Secríst H;
XX WPI: 2000-638466/61.
XX DR N-PSDB; C79150.
XX PT Novel lung tumor polypeptides and polynucleotides, useful for
XX detecting, monitoring or treating cancer, especially lung cancer -
XX Claim 1; Page 171-172; 243pp; English.
XX CC The present sequence is given in a specification relating to compounds
XX for therapy and diagnosis of lung cancer. Polypeptides comprising at
XX least an immunogenic part of a lung tumour protein are disclosed.
XX The polypeptides are useful for inhibiting the development of cancer,
XX especially lung cancer. Samples of T cells expressing the polypeptides
XX may be used to inhibit the development of cancer. The polypeptides are
XX also useful for detecting and monitoring the progression of cancer,
XX especially lung cancer.
XX SQ Sequence 704 AA;

Query Match 50.6%; Score 40.5; DB 21; Length 704;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

QY 1 QLIDIVDQLKNYVNDL 16
::|::|::|::|::|
Db 296 ritdvlqg-knyveel 310

RESULT 10
Y35668
ID Y35668 standard; Protein; 171 AA.
XX AC Y35668;
XX DT 13-SEP-1999 (first entry)
XX DE Chlamydia pneumoniae transport polypeptide.
XX KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
XX sinusitis; purulent otitis media; erythema nodosum; pharyngitis;

```

KW vaccine; neutralising epitope.
 XX Chlamydia pneumoniae.
 OS WO9927105-A2.
 XX
 PN 03-JUN-1999.
 XX
 PD 20-NOV-1998; 98WO-IB01890.
 XX
 PF 04-NOV-1998; 98US-0107078.
 XX
 PR 21-NOV-1997; 97FR-0014673.
 XX
 PA (GEST) GENSET.
 XX
 PI Griffais R;
 XX
 DR WPI; 1999-357842/30.
 XX
 PF Genome sequence of Chlamydia pneumoniae
 XX
 PS Page 1386; Disclosure; 1912pp; English.
 XX
 CC Y34584-Y35879 represent the proteins encoded by all the open reading
 CC frames in the complete genome (see X91990) of Chlamydia pneumoniae.
 CC C. pneumoniae causes respiratory disease such as pneumonia and
 CC bronchitis and is thought to be a contributing factor in heart
 CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
 CC nodosum or pharyngitis. The polypeptides encoded by the open reading
 CC frames of the C. pneumoniae genome (see Y34584-Y35879) can be used in
 CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
 CC nucleotide sequences can also be used as immunogenic compositions,
 CC especially where the vector directs the expression of a neutralising
 CC epitope of C. pneumoniae.
 CC note: the amino terminal of this protein does not appear in the
 CC specification. The actual length of the protein is 443 amino
 CC acids, but only the last 171 amino acids are given.
 XX
 SQ Sequence 171 AA;

 Query Match 50.0%; Score 40; DB 20; Length 171;
 Best Local Similarity 58.3%; Pred. No. 61;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

 QY 5 IVDQLKNYVNDL 16
 Db |::|||::||
 67 iadrlknllndi 78

 RESULT 11
 Y95040
 ID Y95040 standard; Protein; 478 AA.
 XX
 AC Y95040;
 XX
 DT 23-JUN-2000 (first entry)
 XX
 DE Candida albicans polypeptide sequence # 8.
 XX
 KW Candida albicans infection; growth; survival; medicament; AIDS;
 KW vulvovaginitis; immunocompromised patient; treat.
 XX
 OS Candida albicans.
 XX
 PN EP982401-A2.
 XX
 PD 01-MAR-2000.
 XX
 PF 23-DEC-1998; 98EP-0310694.
 XX
 PR 14-AUG-1998; 98GB-0017796.
 XX
 XX

 PA (JANC) JANSSEN PHARM NV.
 XX
 PI Contreras RH, Nelissen B, De Backer MD, Luyten WHML, Viaene JE;
 PI Logghe MG;
 XX
 DR WPI; 2000-258614/23.
 XX
 PT Essential polypeptides isolated from Candida albicans, useful in the
 PT treatment of diseases caused by C.albicans, especially in
 PT immunocompromised subjects, e.g., AIDS patients
 XX
 PS Claim 3; Page 35-36; 133pp; English.
 XX
 CC This sequence represents a polypeptide that is critical for the survival
 CC and growth of Candida albicans. The C. albicans nucleic acid molecules
 CC encoding the polypeptides of the invention may be used as probes and
 CC primers for detecting homologous nucleic acid molecule sequences. The
 CC polypeptides and nucleic acid molecules and compounds identified as
 CC selectively modulating the expression of the polypeptides, may be used as
 CC medicaments or for the preparation of a medicament to treat C.albicans
 CC associated diseases, especially in AIDS patients and to treat
 CC vulvovaginitis in otherwise healthy females. The use of the polypeptides
 CC and polynucleotide sequences to treat C.albicans associated diseases has
 CC fewer side effects and less toxicity than previously used methods such as
 CC the use of amphotericin. This method is therefore especially suitable for
 CC immunocompromised patients, such as AIDS patients.
 XX
 SQ Sequence 478 AA;

 Query Match 50.0%; Score 40; DB 21; Length 478;
 Best Local Similarity 58.3%; Pred. No. 1.9e+02;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

 QY 4 DIVDQLKNYVND 15
 Db |::|||::||
 282 dvidqdlnivdd 293

 RESULT 12
 W81514
 ID W81514 standard; Protein; 541 AA.
 XX
 AC W81514;
 XX
 DT 02-FEB-1999 (first entry)
 XX
 DE Chromodomain Y b (CDYb) gene product.
 XX
 KW Non-recombining region; human; Y chromosome; X homologue; testis; CDYb;
 KW infertility; sperm; inhibitor; gene alteration; chromodomain Y.
 XX
 OS Homo sapiens.
 XX
 PN W09846747-A2.
 XX
 PD 22-OCT-1998.
 XX
 PF 10-APR-1998; 98WO-US07115.
 XX
 PR 11-APR-1997; 97US-0041877.
 XX
 PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
 XX
 PI Lahn BT, Page DC;
 XX
 DR WPI; 1998-568729/48.
 DR N-PSDB; V69644.
 XX
 PT Novel genes in the non-combining region of Y chromosome - useful to
 PT diagnose if male infertility or reduced sperm count has a genetic
 PT basis
 XX

PS Claim 3; Fig 8; 54pp; English.

CC This is the amino acid sequence of Chromodomain Y b (CDYb) gene product.
 CC The invention relates to genes occurring on the non-recombining region
 CC of the human Y chromosome. The sequences fall into two classes: (1)
 CC X-homologous DNA which are expressed in many organs, having functional
 CC X-homologues and (2) testis-specific DNA sequences. Y chromosomal DNA
 CC from males with known conditions such as infertility and reduced sperm
 CC count can be assessed using the invention to determine whether the
 CC condition is associated with or caused by the occurrence of the gene or
 CC gene alteration. Candidate inhibitors of the enzymatic activity of the
 CC genes can be assessed using in vitro assays.

XX Sequence 541 AA;

Query Match 50.0%; Score 40; DB 19; Length 541;

Best Local Similarity 50.0%; Pred. No. 2.2e+02;

Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 3 IDIVDQLKNYVN 14

Db 364 lemvdtiknfvn 375

RESULT 13

W81513
 ID W81513 standard; Protein; 554 AA.

AC W81513;

DT 02-FEB-1999 (first entry)

XX Chromodomain Y a (CDYa) gene product.

XX Non-recombining region; human; Y chromosome; X homologue; testis; CDYa;
 KW infertility; sperm; inhibitor; gene alteration; chromodomain Y.

XX Homo sapiens.

XX WO9846747-A2.

XX 22-OCT-1998.

XX 10-APR-1998; 98WO-0507115.

XX 11-APR-1997; 97US-0041877.

XX (WHED) WHITEHEAD INST BIOMEDICAL RES.

XX Lahn BT, Page DC;

XX WPI: 1998-568729/48.

XX N-PSDB; V69645.

XX Novel genes in the non-combining region of Y chromosome - useful to
 PT diagnose if male infertility or reduced sperm count has a genetic
 PT basis

PS Claim 3; Fig 8; 54pp; English.

CC This is the amino acid sequence of Chromodomain Y a (CDYa) gene product.
 CC The invention relates to genes occurring on the non-recombining region
 CC of the human Y chromosome. The sequences fall into two classes: (1)
 CC X-homologous DNA which are expressed in many organs, having functional
 CC X-homologues and (2) testis-specific DNA sequences. Y chromosomal DNA
 CC from males with known conditions such as infertility and reduced sperm
 CC count can be assessed using the invention to determine whether the
 CC condition is associated with or caused by the occurrence of the gene or
 CC gene alteration. Candidate inhibitors of the enzymatic activity of the
 CC genes can be assessed using in vitro assays.

XX Sequence 554 AA;

Query Match 50.0%; Score 40; DB 19; Length 554;

Best Local Similarity 50.0%; Pred. No. 2.2e+02;

Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 3 IDIVDQLKNYVN 14

Db 363 lemvdtiknfvn 374

RESULT 14

R92289
 ID R92289 standard; Protein; 934 AA.

XX R92289;

DT 18-MAY-1996 (first entry)

XX Cycloisomaltoligosaccharide synthase mature enzyme.

XX Cycloisomaltoligosaccharide synthase; CIS; inclusion compound;
 KW dextran.

XX Bacillus sp. strain T-3040 (FERM BP-4132).

XX EP699749-A1.

XX 06-MAR-1996.

XX 30-AUG-1995; 95EP-0113645.

XX 30-AUG-1994; 94JP-0205631.

XX (KIKK) KIKKOMAN CORP.

XX (NODA) NODA INST SCI RES.

XX (NODA) ZH NODA SANGYO KAGAKU KENKYUSHO.

XX Kurokawa T, Oguma T, Tobe K;

XX WPI: 1996-130767/14.

XX DNA encoding cyclo:iso:malto:oligosaccharide synthase - used to
 PT produce cyclo:iso:malto:oligosaccharide for use in inclusion cpds.

XX Claim 1; Page 14-18; 20pp; English.

XX Cycloisomaltoligosaccharide synthase (CIS) (R92289) was isolated
 CC from the culture broth of a soil isolate, Bacillus sp. T-3040. The
 CC enzyme is capable of converting alpha-1,6-glucans, esp. dextran,
 CC into cycloisomaltoligosaccharides (CIOS). The CIOS form inclusion
 CC cpds. useful in pharmaceutical prepn., food, etc., and as
 CC anticaries agents. A gene (see T15937) coding for the CIS precursor
 CC polypeptide (R92288) has been isolated and can be used for
 CC efficient prodn. of the enzyme, pref. in Escherichia coli host cell
 CC transformants.

XX Sequence 934 AA;

Query Match 50.0%; Score 40; DB 17; Length 934;

Best Local Similarity 53.3%; Pred. No. 3.9e+02;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 2 LIDIVDQLKNYVNDL 16

Db 346 ldsynqknknyieql 360

RESULT 15

R92288
 ID R92288 standard; Protein; 972 AA.

XX

AC R92288; /
XX
XX
DT 18-MAY-1996 (first entry)
XX
DE Cycloisomaltoligosaccharide synthase precursor.
XX
XX Cycloisomaltoligosaccharide synthase; CIS; inclusion compound;
KW dextran.
XX
XX
OS Bacillus sp. strain T-3040 (FERM BP-4132).
XX
XX
FH Key Location/Qualifiers
FT Peptide 1..38
FT /label= Sig_peptide
XX
XX EP699749-A1.
XX
XX
PD 06-MAR-1996.
XX
XX 30-AUG-1995; 95EP-0113645.
XX
XX 30-AUG-1994; 94JP-0205631.
XX
XX (KIKK) KIKKOMAN CORP.
PA (NODA) NODA INST SCI RES.
PA (NODA) ZH NODA SANGYO KAGAKU KENKYUSHO.
XX
XX Kurokawa T, Oguma T, Tobe K;
PI WPI; 1996-130767/14.
XX
XX N-PSDB; T15957.
XX
XX
PT DNA encoding cyclo:iso:malto:oligosaccharide synthase - used to
PT produce cyclo:iso:malto:oligosaccharide for use in inclusion cpds.
XX
XX
PS Claim 1; Page 7-11; 20pp; English.
XX
XX Cycloisomaltoligosaccharide synthase (CIS) (R92288) is encoded by
CC a gene (T15957) isolated from a genomic DNA library of a soil
CC isolate, Bacillus sp. T-3040. CIS can be obtd. efficiently in short
CC culture times by expression of the gene in a prokaryotic or
CC eukaryotic recombinant host cell, pref. Escherichia coli. The
CC mature enzyme (see also R92289) is used for the prodn. of
CC cycloisomaltoligosaccharides (CIos) from alpha-1,6-glucans, pref.
CC dextran. The CIos form inclusion cpds. useful in pharmaceutical
CC preps., food, etc., and as anticaries agents.
XX
XX
SQ Sequence 972 AA;

Query Match 50.0%; Score 40; DB 17; Length 972;
Best Local Similarity 53.3%; Pred. NO. 4.1e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 LIDIVDQLKKNVNDL 16
| | :||||| :
Db 384 lsdynqlknyieql 398

Search completed: May 23, 2001, 11:11:35
Job time: 178 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:09:42 ; Search time 58.85 Seconds
(without alignments)
5.223 Million cell updates/sec

Title: US-09-522-217-2_COPY_41_56

Perfect score: 80

Sequence: 1 QLIDIVDQLKNVNDL 16

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2.6/prodata/2/iaa/5A_COMB.pep.*

2: /cgn2.6/prodata/2/iaa/5B_COMB.pep.*

3: /cgn2.6/prodata/2/iaa/6A_COMB.pep.*

4: /cgn2.6/prodata/2/iaa/6B_COMB.pep.*

5: /cgn2.6/prodata/2/iaa/PCTUS_COMB.pep.*

6: /cgn2.6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	51.2	830	3 US-08-804-439A-20	Sequence 20, Appl
2	41	51.2	830	3 US-08-720-229-20	Sequence 20, Appl
3	40	50.0	309	3 US-09-058-489-7	Sequence 7, Appl
4	40	50.0	541	3 US-09-058-489-40	Sequence 40, Appl
5	40	50.0	554	3 US-09-058-489-39	Sequence 39, Appl
6	39	48.8	377	2 US-08-839-581A-31	Sequence 31, Appl
7	39	48.8	377	4 US-09-023-591A-31	Sequence 31, Appl
8	39	48.8	805	1 US-08-045-806-2	Sequence 2, Appl
9	39	48.8	805	1 US-08-366-051B-2	Sequence 2, Appl
10	38	47.5	156	3 US-08-765-381-11	Sequence 11, Appl
11	38	47.5	572	3 US-09-040-681A-4	Sequence 4, Appl
12	37	46.2	203	4 US-09-124-141-9	Sequence 9, Appl
13	37	46.2	250	4 US-09-124-141-17	Sequence 17, Appl
14	37	46.2	593	4 US-09-124-141-7	Sequence 7, Appl
15	37	46.2	594	2 US-08-592-696-2	Sequence 2, Appl
16	37	46.2	594	2 US-08-592-696-4	Sequence 4, Appl
17	37	46.2	594	2 US-09-027-536-2	Sequence 2, Appl
18	37	46.2	594	2 US-09-027-536-4	Sequence 4, Appl
19	37	46.2	594	3 US-09-028-148-2	Sequence 2, Appl
20	37	46.2	594	3 US-09-028-148-4	Sequence 4, Appl
21	37	46.2	594	4 US-09-124-141-15	Sequence 15, Appl
22	37	46.2	594	4 US-09-124-141-23	Sequence 23, Appl
23	37	46.2	625	1 US-08-242-689-2	Sequence 2, Appl
24	37	46.2	848	1 US-08-045-806-4	Sequence 4, Appl
25	37	46.2	848	1 US-08-366-051B-4	Sequence 4, Appl
26	37	46.2	953	2 US-08-506-340A-1	Sequence 1, Appl
27	37	46.2	1088	2 US-08-742-026-2	Sequence 2, Appl

28 37 46.2 1088 2 US-08-742-026-23 Sequence 23, Appl
29 36 45.0 20 3 US-08-718-905-5 Sequence 5, Appl
30 36 45.0 141 1 US-08-284-393B-10 Sequence 10, Appl
31 36 45.0 141 1 US-08-259-696B-11 Sequence 11, Appl
32 36 45.0 141 2 US-08-902-513-11 Sequence 11, Appl
33 36 45.0 141 5 PCT-US95-08950-10 Sequence 10, Appl
34 36 45.0 145 1 US-08-470-179-32 Sequence 32, Appl
35 36 45.0 164 1 US-08-318-193-77 Sequence 77, Appl
36 36 45.0 165 2 US-08-955-848A-1 Sequence 1, Appl
37 36 45.0 165 5 PCT-US95-03866-2 Sequence 2, Appl
38 36 45.0 166 2 US-08-628-428-2 Sequence 5, Appl
39 36 45.0 166 2 US-08-628-428-5 Sequence 8, Appl
40 36 45.0 166 2 US-08-628-428-8 Sequence 2, Appl
41 36 45.0 166 2 US-09-106-891-2 Sequence 18, Appl
42 36 45.0 166 5 PCT-US95-03866-18 Sequence 34, Appl
43 36 45.0 166 5 PCT-US95-03866-34 Sequence 36, Appl
44 36 45.0 166 5 PCT-US95-03866-36 Sequence 20, Appl
45 36 45.0 167 5 PCT-US95-03866-20

ALIGNMENTS

RESULT 1
US-08-804-439A-20
; Sequence 20, Application US/08804439A
; Patent No. 6015565
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Ste 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,439A
; FILING DATE: February 21, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09176/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 830 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-804-439A-20

Query Match 51.2%; Score 41; DB 3; Length 830;
Best Local Similarity 53.3%; Pred. No. 1.2e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QLIDIVDQLKNVND 15

Query Match 50.0%; Score 40; DB 3; Length 554;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 3 IDIVIDQLKNYVN 14

Db 363 LEMVDTIKNEVN 374

RESULT 6

US-08-839-581A-31

; Sequence 31, Application US/08839581A

; Patent No. 5958705

; GENERAL INFORMATION:

; APPLICANT: Staunton, Donald E.

; APPLICANT: Lipsky, Brian P.

; TITLE OF INVENTION: Cytoplasmic Modulators of Integrin

; TITLE OF INVENTION: Binding/Signaling

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 233 South Wacker, 6300 Sears Tower

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/839,581A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Young J. Suh

; REGISTRATION NUMBER: P-41,337

; REFERENCE/DOCKET NUMBER: 27866/33886

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-474-6300

; TELEFAX: 312-474-0448

; INFORMATION FOR SEQ ID NO: 31:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 377 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-839-581A-31

Query Match 48.8%; Score 39; DB 2; Length 377;
Best Local Similarity 43.8%; Pred. No. 1e+02;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 QLIDIVDLKNYVNDL 16

Db 12 QLVDEIEALKNEVREV 27

RESULT 7

US-09-023-591A-31

; Sequence 31, Application US/09023591A

; Patent No. 6210914

; GENERAL INFORMATION:

; APPLICANT: Staunton, Donald E.

; APPLICANT: Lipsky, Brian P.

; TITLE OF INVENTION: Cytoplasmic Modulators of Integrin

; TITLE OF INVENTION: Binding/Signaling

; NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/023,591A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/839,581

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Young J. Suh

REGISTRATION NUMBER: P-41,337

REFERENCE/DOCKET NUMBER: 27866/33886

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300

TELEFAX: 312-474-0448

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:

LENGTH: 377 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-023-591A-31

Query Match 48.8%; Score 39; DB 4; Length 377;

Best Local Similarity 43.8%; Pred. No. 1e+02;

Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 QLIDIVDLKNYVNDL 16

Db 12 QLVDEIEALKNEVREV 27

RESULT 8

US-08-045-806-2

; Sequence 2, Application US/08045806

; Patent No. 5378822

; GENERAL INFORMATION:

; APPLICANT: Bradfield, Christopher Alan

; APPLICANT: Dolwick, Kristin Marie

; APPLICANT: Poland, Alan

; TITLE OF INVENTION: Ah Receptor cDNA and Method of

; TITLE OF INVENTION: Determining Human Risks To Environmental Pollutants

; NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Tilton, Fallon, Lungmus & Chestnut

STREET: 100 South Wacker Drive, Suite 960

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606-4002

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/045,806

FILING DATE: 19930408

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

```

; NAME: Fentress, Susan B.
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: NU-9207
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)-456-8000
; TELEFAX: (312)-456-7776
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 805 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-045-806-2

Query Match 48.8%; Score 39; DB 1; Length 805;
Best Local Similarity 53.8%; Pred. No. 2.3e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 IDIVDLKKNVND 15
Db 569 IDITDEILTVQD 581

RESULT 9
US-08-366-051B-2
; Sequence 2, Application US/08366051B
; Patent No. 5650283
; GENERAL INFORMATION:
; APPLICANT: Bradfield, Christopher A.
; APPLICANT: Dolwick, Kristin M.
; APPLICANT: Carver, Lucy A.
; TITLE OF INVENTION: Ah Receptor cDNAs and Genetically
; TITLE OF INVENTION: Engineered Cells for Detecting Agonists to the Ah
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Tilton, Fallon, Lungmus & Chestnut
; STREET: 100 South Wacker Drive, Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/366,051B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Tilton, Timothy L.
; REGISTRATION NUMBER: 16,926
; REFERENCE/DOCKET NUMBER: NU-9207-CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)-456-8000
; TELEFAX: (312)-456-7776
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 805 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-366-051B-2

Query Match 48.8%; Score 39; DB 1; Length 805;
Best Local Similarity 53.8%; Pred. No. 2.3e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 IDIVDLKKNVND 15

```

ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040.681A
FILING DATE: 18 MAR 98
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/041,300
FILING DATE: 18 MAR 97
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 487900-16 (N)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 572 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-09-040-681A-4

```
Query Match      47.5%; Score 38; DB 3; Length 572;
Best Local Similarity 40.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      2 LIDIVDQLKNYVNDL 16
      |||:||||:::  :|:
Db     374 LIDVDEIRSKRSDI 388
```

```

RESULT 12
US-09-124-141-9
; Sequence 9, Application US/09124141
; Patent No. 6211352
; GENERAL INFORMATION:
; APPLICANT: Harrison, Leonard
; APPLICANT: Honeyman, Margot
; APPLICANT: Cram, David
; APPLICANT: De Aizpurua, Henry
; TITLE OF INVENTION: A METHOD FOR THE DIAGNOSIS AND TREATMENT OF GLUTAMIC
; TITLE OF INVENTION: ACID DECARBOXYLASE AUTOANTIGEN ASSOCIATED DISEASES
; FILE REFERENCE: Phillips, Ormonde & Fitzpatrick
; CURRENT APPLICATION NUMBER: US/09/124,141
; CURRENT FILING DATE: 1998-07-29
; EARLIER APPLICATION NUMBER: 08/308,952
; EARLIER FILING DATE: 1994-09-20
; EARLIER APPLICATION NUMBER: 07/839,805
; EARLIER FILING DATE: 1992-02-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: N-terminal
; OTHER INFORMATION: fragment of mouse brain GAD (MBGAD12)
US-09-124-141-9

```

Query Match 46.2%; Score 37; DB 4; Length 203;
Best Local Similarity 58.3%; Pred. No. 1e+02;
Matches 7; Conservative 3; Mismatches 2; Indels

Qy 2 LIDIVDLKNYV 13
|::|||
Db 116 LLEVDPILLNYV 127

RESULT 13
US-09-124-141-17
; Sequence 17, Application US/09124141
; Patent No. 6211352
; GENERAL INFORMATION:
; APPLICANT: Harrison, Leonard
; APPLICANT: Honeyman, Margot
; APPLICANT: Cram, David
; APPLICANT: De Aizourua, Henry
; TITLE OF INVENTION: A METHOD FOR THE DIAGNOSIS AND TREATMENT OF GLUTAMIC
; TITLE OF INVENTION: ACID DECARBOXYLASE AUTOANTIGEN ASSOCIATED DISEASES
; FILE REFERENCE: Phillips, Ormonde & Fitzpatrick
; CURRENT APPLICATION NUMBER: US/09/124,141
; CURRENT FILING DATE: 1998-07-29
; EARLIER APPLICATION NUMBER: 08/308,952
; EARLIER FILING DATE: 1994-09-20
; EARLIER APPLICATION NUMBER: 07/839,805
; EARLIER FILING DATE: 1992-02-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 250
; TYPE: prt
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:N-terminal
; OTHER INFORMATION: Fragment of Human Brain GAD (HBGAD17)
US-09-124-141-17

Query Match	46.2%;	Score 37;	DB 4;	Length 250;
Best Local Similarity	58.3%;	Pred. No. 1.3e+02;		
Matches	7;	Conservative	3;	Mismatches 2;
				Indels 0;
				Gaps 0;

QY 2 LIDIVDLKNYV 13
|::||| | |||
Db 117 LLEVVDILLNYV 128

RESULT 14
US-09-124-141-7
; Sequence 7, Application US/09124141
; Patent No. 6211352
; GENERAL INFORMATION:
; APPLICANT: Harrison, Leonard
; APPLICANT: Honeyman, Margot
; APPLICANT: Cram, David
; APPLICANT: De Aizpuru, Henry
; TITLE OF INVENTION: A METHOD FOR THE DIAGNOSIS AND TREATMENT OF GLUTAMIC
; TITLE OF INVENTION: ACID DECARBOXYLASE AUTOANTIGEN ASSOCIATED DISEASES
; FILE REFERENCE: Phillips, Ormonde & Fitzpatrick
; CURRENT APPLICATION NUMBER: US/09/124,141
; CURRENT FILING DATE: 1998-07-29
; EARLIER APPLICATION NUMBER: 08/308,952
; EARLIER FILING DATE: 1994-09-20
; EARLIER APPLICATION NUMBER: 07/839,805
; EARLIER FILING DATE: 1992-02-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 593
; TYPE: prt
; ORGANISM: Unknown Organism

; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Full Length Mouse
; OTHER INFORMATION: Brain GAD
US-09-124-141-7

Query Match 46.2%; Score 37; DB 4; Length 593;
Best Local Similarity 58.3%; Pred. No. 3.3e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LIDIVDQLKNYV 13
|::|||
Db 116 LLEVVDILLNYV 127

RESULT 15
US-08-592-696-2
; Sequence 2, Application US/08592696
; Patent No. 5821334
; GENERAL INFORMATION:
; APPLICANT: Powers, Alvin C
; TITLE OF INVENTION: "INSULIN-DEPENDENT DIABETES
; TITLE OF INVENTION: MELLITUS-SPECIFIC CHIMERIC POLYPEPTIDES"
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Needle & Rosenberg, P.C.
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592.696
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Selby Esq., Elizabeth
; REGISTRATION NUMBER: 36,298
; REFERENCE/DOCKET NUMBER: 22000.0043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 594 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-592-696-2

Query Match 46.2%; Score 37; DB 2; Length 594;
Best Local Similarity 58.3%; Pred. No. 3.3e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LIDIVDQLKNYV 13
|::|||
Db 117 LLEVVDILLNYV 128

Search completed: May 23, 2001, 11:09:43
Job time: 66 sec

GenCore version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:12:51 ; Search time 70.54 Seconds
(without alignments)
15.588 Million cell updates/sec

Title: US-09-522-217-2_COPY_41_56
Perfect score: 80
Sequence: 1 QLIDIVDQLKNVNDL 16

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_67: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	57.5	569	2	S41200
2	45	56.2	264	1	phosphoglucumutase
3	44	55.0	213	2	indole-3-glycerol-
4	44	55.0	304	2	hypothetical prote
5	44	55.0	471	2	translation initia
6	44	55.0	935	2	hypothetical prote
7	43	53.8	189	2	hypothetical prote
8	43	53.8	325	2	probable membrane
9	43	53.8	521	2	hypothetical prote
10	43	53.8	608	2	hypothetical prote
11	42	52.5	235	2	ORF MSV140 hypothe
12	42	52.5	286	1	hypothetical prote
13	42	52.5	358	2	thymidylate synth
14	42	52.5	559	2	N5, N10-methylenete
15	42	52.5	1397	2	DNA primase - Heli
16	41	51.2	78	2	DNA topoisomerase
17	41	51.2	183	2	hypothetical prote
18	41	51.2	243	2	hypothetical prote
19	41	51.2	412	2	cytolysin activato
20	41	51.2	462	2	L1al restriction a
21	41	51.2	518	2	dihydrofolate redu
22	41	51.2	552	2	Lztrr-1 - human
23	41	51.2	649	2	hypothetical prote
24	41	51.2	669	2	ATP-dependent DNA
25	41	51.2	724	2	probable membrane
26	41	51.2	769	2	hypothetical prote
27	41	51.2	809	2	UDP-N-acetylmuram
28	41	51.2	830	1	glycoprotein B pre
29	41	51.2	830	1	glycoprotein B pre

30	41	51.2	830	2	T43999	glycoprotein B [im
31	41	51.2	830	2	T44186	probable glycoprot
32	41	51.2	1279	2	T41389	rna binding protei
33	40.5	50.6	430	2	T20170	hypothetical prote
34	40	50.0	69	2	D82933	type I restriction
35	40	50.0	72	2	E69386	hypothetical prote
36	40	50.0	150	2	E42528	B24k protein - vac
37	40	50.0	206	2	T29985	hypothetical prote
38	40	50.0	359	2	G72425	hypothetical prote
39	40	50.0	370	2	T28376	ORF MSV215 SCG gen
40	40	50.0	445	2	C71636	probable thiophene
41	40	50.0	528	2	S35272	dihydrofolate redu
42	40	50.0	565	2	D72006	sulfate transporte
43	40	50.0	577	2	T14315	dihydrofolate redu
44	40	50.0	831	1	VBGE6S	glycoprotein B - h
45	40	50.0	900	2	C64232	alanine--CRNA liga

ALIGNMENTS

RESULT 1

S41200
phosphoglucumutase (EC 5.4.2.2) PGM2 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YW9718.04c; protein YMR105c
C:Species: Saccharomyces cerevisiae
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 16-Jul-1999
C:Accession: S41200; S54566; S40264
R:Boles, E.; Liebetrau, W.; Hofmann, M.; Zimmermann, F.K.
Eur. J. Biochem. 220, 83-96, 1994
A:Title: A family of hexosephosphate mutases in Saccharomyces cerevisiae.
A:Reference number: S41199; MUID:94164176
A:Accession: S41200
A:Molecule type: DNA
A:Residues: 1-569 <BOL>
A:Cross-references: EMBL:X74823; NID:9438233; PIDN:CAA52820.1; PID:9438234
R:Hunt, S.; Bowman, S.
submitted to the EMBL Data Library, May 1995
A:Reference number: S54510
A:Accession: S54566
A:Molecule type: DNA
A:Residues: 1-569 <HUN>
A:Cross-references: EMBL:Z49702; NID:9817859; PIDN:CAA89741.1; PID:9817863; MIPS:YMR1
A:Experimental source: strain AB972
C:Genetics:
A:Gene: SGD:PGM2
A:Cross-references: SGD:S0004711; MIPS:YMR105c
A:Map position: 13R
C:Superfamily: phosphoglucumutase
C:Keywords: intramolecular transferase; isomerase

Query Match 57.5%; Score 46; DB 2; Length 569;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 LIDIVDQLKNVNDL 16
DB 183 LVDIIDTKDIVNFL 197

RESULT 2

S35127
indole-3-glycerol-phosphate synthase (EC 4.1.1.48) - Lactococcus lactis subsp. lactis
C:Species: Lactococcus lactis subsp. lactis
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S35127
R:Bardowski, J.; Ehrlich, S.D.; Chopin, A.
J. Bacteriol. 174, 6563-6570, 1992
A:Title: tryptophan biosynthesis genes in Lactococcus lactis subsp. lactis.
A:Reference number: S35123; MUID:93015708
A:Accession: S35127
A:Molecule type: DNA

A:Residues: 1-264 <BAR>
A:Cross-references: EMBL:M87483; NID:gl49514; PIDN:AAA25226.1; PID:gl49519
C:Genetics:
A:Gene: trpC
C:Superfamily: indole-3-glycerol-phosphate synthase; trpC homology
C:Keywords: carbon-carbon lyase; carboxy-lyase; tryptophan biosynthesis
F:8-255/Domain: trpC homology <TRC>

Query Match 56.2%; Score 45; DB 1; Length 264;
Best Local Similarity 80.0%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 IDIVIDQKKNY 12
 :|||||
Db 68 VDIVDQAKNY 77

RESULT 3
T27898
hypothetical protein ZK546.11 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T27898
R:Hallsworth, K.
Submitted to the EMBL Data Library, June 1995
A:Description: The sequence of *C. elegans* cosmid ZK546.
A:Reference number: Z20437
A:Accession: T27898
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-213 <HAL>
A:Cross-references: EMBL:U29380; PIDN:AAA68744.1; CESP:ZK546.11
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:ZK546.11
A:Introns: 45/1
C:Superfamily: glutathione transferase

Query Match 55.0%; Score 44; DB 2; Length 213;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 QLIDIVDLKKNYVNDL 16
 : |||||
Db 87 QVNAIDQFKDYNDI 102

RESULT 4
A32108
translation initiation factor eIF-2 alpha chain - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein J1429; protein YJR007w
C:Species: *Saccharomyces cerevisiae*
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 26-Aug-1999
C:Accession: A32108; S55195; S57022
R:Cigan, A.M.; Pabich, E.K.; Feng, L.; Donahue, T.F.
Proc. Natl. Acad. Sci. U.S.A. 86, 2784-2788, 1989
A:Title: Yeast translation initiation suppressor su12 encodes the alpha-subunit of eukaryotic translation initiation factor eIF-2
A:Reference number: A32108; MUID:89202411
A:Accession: A32108
A:Molecule type: DNA
A:Residues: 1-304 <CIG>
A:Cross-references: EMBL:M25552; NID:g341369; PIDN:AAA70332.1; PID:g903889
R:de Haan, M.; Smits, P.H.M.; Grivell, L.A.
Submitted to the EMBL Data Library, May 1995
A:Reference number: S55183
A:Accession: S55195
A:Molecule type: DNA
A:Residues: 1-304 <DBH>
A:Cross-references: EMBL:X87611; NID:g854567; PIDN:CAA60929.1; PID:g854580
R:de Haan, M.; Grivell, L.A.; Smits, P.H.M.
Submitted to the Protein Sequence Database, September 1995

A:Reference number: S56771
A:Accession: S57022
A:Molecule type: DNA
A:Residues: 1-304 <ZAG>
A:Cross-references: EMBL:Z49507; NID:gl015631; PIDN:CAA89529.1; PID:gl015632; MIPS:YU
C:Genetics:
A:Gene: SGD:SUI2
A:Cross-references: SGD:S0003767; MIPS:YJR007w
A:Map position: 10R
C:Superfamily: translation initiation factor eIF-2 alpha chain
C:Keywords: phosphoprotein; protein biosynthesis

Query Match 55.0%; Score 44; DB 2; Length 304;
Best Local Similarity 54.5%; Pred. No. 29;
Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 4 DIVDQLKNYVN 14
 |::|:|::|
Db 164 DVLDELKNYIS 174

RESULT 5
E64223
hypothetical protein homolog MG213 - *Mycoplasma genitalium*
C:Species: *Mycoplasma genitalium*
C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 07-Dec-1999
C:Accession: E64223
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.
M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, C.A.; Venter, J.C.
Science 270, 397-403, 1995
A:Title: The minimal gene complement of *Mycoplasma genitalium*.
A:Reference number: A64200; MUID:96026346
A:Accession: E64223
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-471 <TIGR>
A:Cross-references: GB:U39698; GB:L43967; NID:gl045891; PID:gl045899; TIGR:MG213
A:Experimental source: strain G-37
C:Genetics:
A:Genetic code: SGC3
C:Superfamily: hypothetical protein homolog MG213; type I dihydrofolate reductase hom

Query Match 55.0%; Score 44; DB 2; Length 471;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 QLIDIVDLKKNYVNDL 16
 |::|:|::|
Db 288 QLVTVEDQLPNYISVL 303

RESULT 6
S57080
hypothetical protein YJR061w - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: hypothetical protein J1736
C:Species: *Saccharomyces cerevisiae*
C:Date: 08-Sep-1995 #sequence_revision 08-Sep-1995 #text_change 29-Oct-1999
R:Huang, M.E.; Chuat, J.C.; Galibert, F.
Submitted to the Protein Sequence Database, September 1995
A:Reference number: S57052
A:Accession: S57080
A:Molecule type: DNA
A:Residues: 1-935 <MAN>
A:Cross-references: EMBL:Z49561; NID:gl015732; PIDN:CAA89589.1; PID:gl015733; MIPS:YU
R:Baker, R.T.; Varshavsky, A.
Submitted to the EMBL Data Library, September 1994
A:Description: N-terminal amidase: a new enzyme and component of a targeting complex
A:Reference number: S47937
A:Accession: S47939

Genetics:
Map position: 10R

Query Match 53.8%; Score 43; DB 2; Length 325;
Best Local Similarity 58.3%; Pred. NO. 44;
Matches 7: Conservative 4; Mismatches 1; Indels

QY 4 DIVDQLKNYVND 15
: : : : :
Db 283 EVVDVLKSYLND 294

RESULT 9
T01684
dihydrofolate reductase (EC 1.5.1.3) / thymidylate synthase (EC 2.1.1.45) - maize
C:Species: Zea mays (maize)
C:date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 23-Jul-1999
C:Accession: T01684

Query Match 53.8%; Score 43; DB 2; Length 521;
Best Local Similarity 46.7%; Pred. No. 70;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0

Query Match 53.8%; Score 43; DB 2; Length 521;
Best Local Similarity 46.7%; Pred. No. 70;
Matches 7; Conservative 5; Mismatches 3; Indels

QY 1 QLIDIVDLKNYVND 15
||:|:|:|:| 1
Db 368 QLMVDIKIKNDPED 382

RESULT 10
T28301
ORF MSV140 hypothetical protein - Melanoplus sanguinipes entomopoxvirus
C/Species: Melanoplus sanguinipes entomopoxvirus
C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
C/Accession: T28301
R/Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.
J. Virol. 73, 533-552, 1999
A/Title: The genome of Melanoplus sanguinipes entomopoxvirus.
A/Reference number: Z20484; MUID:99102612
A/Accession: T28301
A/Status: preliminary; translated from GB/EMBL/DBDJ
A/Molecule type: DNA
A/Residues: 1-608 <AFO>
A/Cross-references: EMBL:AF063866; NID:g4049647; PIDN:AAC97872.1; PID:g4049912
C/Genetics:
A/Note: MSV140

Query Match 53.8%; Score 43; DB 2; Length 608;

Beet Local Similarity 50.0%; Score 42; DB 2; Length 235;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 3 IDVDQLKNYVNDL 16
| | | : : | | |
Db 316 IDIIEIGNYNDI 329

RESULT 11

T29984
hypothetical protein Fl1G11.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T29984

R:Lafrelle, P.; Deadman, R.

submitted to the EMBL Data Library, November 1996

A:Description: The sequence of C. elegans cosmid Fl1G11.

A:Reference number: 220715

A:Accession: T29984

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-235 <LAT>

A:Cross-references: EMBL:U80451; PIDN:AAB37845.1; GSPDB:GNO00020; CESP:Fl1G11.3

A:Experimental source: strain Bristol N2; clone Fl1G11

C:Genetics:

A:Gene: CESP:Fl1G11.3

A:Map position: 2

A:Introns: 74/1

C:Superfamily: glutathione transferase

Query Match 52.5%; Score 42; DB 2; Length 235;
Best Local Similarity 58.3%; Pred. No. 44;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 5 IDVDQLKNYVNDL 16
| | | : : | | |
Db 120 IADQFKYLVNDV 131

RESULT 12

SYBPT4

thymidylate synthase (EC 2.1.1.45) - phage T4

C:Species: phage T4

A:Note: host Escherichia coli

C:Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 01-Dec-2000

C:Accession: A00550; T10129

R:Chu, F.K.; Maley, G.F.; Maley, F.; Belfort, M.

Proc. Natl. Acad. Sci. U.S.A. 81, 3049-3053, 1984

A:Title: Intervening sequence in the thymidylate synthase gene of bacteriophage T4.

A:Reference number: A00550; MUID:84221902

A:Accession: A00550

A:Molecule type: DNA

A:Residues: 1-286 <CHU>

A:Cross-references: GB:R01804

R:Chu, F.

submitted to the EMBL Data Library, April 1998

A:Reference number: Z16963

A:Accession: T10129

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-31, 'r', 33-286 <CH2>

A:Cross-references: EMBL:M12742; NID:g3033366; PID:g3033367

A:Experimental source: strain alc4

C:Comment: This enzyme is also expressed by the thyA gene of E. coli; the phage and host

C:Genetics:

A:Gene: td

C:Superfamily: thymidylate synthase; thymidylate synthase homology

C:Keywords: deoxyribonucleotide biosynthesis; methyltransferase

F:1-286/Domain: thymidylate synthase homology <DS>

F:156/Active site: Cys #status experimental

Query Match 52.5%; Score 42; DB 1; Length 286;
Best Local Similarity 40.0%; Pred. No. 54;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QLIDIVDQLKNYVND 15
| : | : : | : : | |
Db 121 QIIEVIDRIKKLPND 135

RESULT 13

S30603

N5, N10-methylenetetrahydrodromethanopterin dehydrogenase (H2-forming) (EC 1.12.99.-) - M

C:Species: Methanopyrus kandleri

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 22-Oct-1999

C:Accession: S30603

R:Zirngibl, C.; van Dongen, W.; Schwoerer, B.; von Buerenau, R.; Richter, M.; Klein, A.

Eur. J. Biochem. 208, 511-520, 1992

A:Title: H(2)-forming methylenetetrahydrodromethanopterin dehydrogenase, a novel type of

A:Reference number: S30603; MUID:92394151

A:Accession: S30603

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-358 <ZIR>

A:Cross-references: EMBL:X60719; NID:g44373; PIDN:CAA43127.1; PID:g44374

C:Keywords: oxidoreductase

Query Match 52.5%; Score 42; DB 2; Length 358;
Best Local Similarity 46.2%; Pred. No. 67;
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Qy 4 DIVDQLKNYVNDL 16
| : : | : : | | |
Db 102 DVMPKLRDYVNDI 114

RESULT 14

D64521

DNA primase - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999

C:Accession: D64521

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodok, A.; McKe

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey,

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467

A:Accession: D64521

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-559 <TOM>

A:Cross-references: GB:AE000523; GB:AE000511; NID:g2313077; PIDN:AAD07082.1; PID:g231

C:Superfamily: DNA primase

Query Match 52.5%; Score 42; DB 2; Length 559;
Best Local Similarity 57.1%; Pred. No. 1e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QLIDIVDQLKNYVND 14
| | | | : : | | |
Db 12 QTIIDIVEISSYVN 25

RESULT 15

T10466

DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) II - malaria parasite (Plasmodium f

C:Species: Plasmodium falciparum

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000

C:Accession: T10466

R:Cheesman, S.J.

submitted to the EMBL Data Library, September 1995
 A:Reference number: Z17031
 A:Accession: T10466
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1397 <CHE>
 A:Cross-references: EMBL:X79345; NID:g994807
 C:Genetics:
 A:Gene: TopoII
 A:Map position: 14
 C:Superfamily: eukaryotic type II DNA topoisomerase; phage T4 DNA topoisomerase (ATP-hydrolase)
 C:Keywords: ATP; DNA binding; isomerase; nucleus

Query Match 52.5%; Score 42; DB 2; Length 1397;
 Best Local Similarity 54.5%; Pred. NO. 2.6e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 DIVDOLKNYVN 14
 ||:|:|:|:|
 Db 895 DIIDNIRYIN 905

Search completed: May 23, 2001, 11:12:52
 Job time: 255 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:20:25 ; Search time 40.06 seconds
(without alignments)
13.682 Million cell updates/sec

Title: US-09-522-217-2_COPY_41_56
Perfect score: 80
Sequence: 1 QLIDIVDLKNVNDL 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	47	58.8	519	1 DRTL_ARATH	Q05762 arabidopsis
2	46	57.5	569	1 PG2M_YEAST	P37012 saccharomyc
3	45	56.2	264	1 TRPC_LACLA	Q01999 lactococcus
4	44	55.0	304	1 IF2A_YEAST	P20459 saccharomyc
5	44	55.0	471	1 Y213_MYCGE	P47455 mycoplasma
6	44	55.0	935	1 YJ31_YEAST	P40355 saccharomyc
7	43	53.8	325	1 YJ40_YEAST	P47120 saccharomyc
8	43	53.8	521	1 DRTL_MAIZE	O81395 zea mays (m
9	42	52.5	286	1 TYSI_BPT4	P00471 bacterioph
10	42	52.5	294	1 TYSI_ASCSU	O96650 ascaris suu
11	42	52.5	358	1 HMD_METKA	Q02394 methanopyru
12	42	52.5	559	1 PRIM_HELPY	P56064 helicobacte
13	42	52.5	1398	1 TOP2_PLAFK	P41001 plasmodium
14	41	51.2	78	1 RPSW_THIFE	P24694 thlobacillu
15	41	51.2	518	1 DRTL_ARATH	Q05763 arabidopsis
16	41	51.2	724	1 YMT6_YEAST	Q05029 saccharomyc
17	41	51.2	809	1 MUDD_CHLPN	Q92701 c murc/dcl
18	41	51.2	830	1 VGLB_HSV6G	P16319 human herpe
19	41	51.2	830	1 VGLB_HSV6U	P28864 human herpe
20	41	51.2	830	1 VGLB_HSV6Z	P36320 human herpe
21	40	50.0	112	1 MTRB_METKA	Q32866 methanopyru
22	40	50.0	150	1 VC18_VACCC	P21102 vaccinia vi
23	40	50.0	243	1 TRPA_SULSO	P50382 sulfolobus
24	40	50.0	445	1 THDF_RICPR	Q92cil rickettsia
25	40	50.0	527	1 MET3_CANAL	Q9Y872 candida alb
26	40	50.0	528	1 DRTL_DAUCA	P45350 daucus caro
27	40	50.0	550	1 PM22_LYCES	Q96575 lycopersico
28	40	50.0	900	1 SYA_MYCGE	P47534 mycoplasma
29	40	50.0	972	1 CTAL_BACCI	P94286 bacillus ci
30	40	50.0	2869	1 RBP1_PLAVB	Q00798 plasmodium
31	39.5	49.4	334	1 HRC4_MYCCA	P71498 mycoplasma
32	39	48.8	377	1 YMX4_CAEEL	P34512 caenorhabdi
33	39	48.8	483	1 PRPD_ECOLI	P77243 escherichia

34	39	48.8	483	1 PRPD_SALTY	P74840 salmonella
35	39	48.8	570	1 PG2M_YEAST	P33401 saccharomyc
36	39	48.8	655	1 YKDA_MYCCA	P45615 mycoplasma
37	39	48.8	805	1 AHR_MOUSE	P30561 mus musculu
38	39	48.8	853	1 AHR_RAT	P41738 rattus norv
39	39	48.8	1244	1 Y307_MYCPN	P75342 mycoplasma
40	39	48.8	2291	1 SPCB_DROME	Q00963 drosophila
41	38.5	48.1	588	1 CMC2_CAEEL	Q20799 caenorhabdi
42	38	47.5	63	1 MT3_ACTCH	P43389 actinidia c
43	38	47.5	135	1 YB57_METJA	Q58852 methanococ
44	38	47.5	138	1 FLGN_ECOLI	P43533 escherichia
45	38	47.5	140	1 FLGN_SALTY	P37406 salmonella

ALIGNMENTS

RESULT 1					
DRTL_ARATH					
ID	DRTL_ARATH	STANDARD:	PRT:	519 AA.	
AC	Q05762;				
DT	01-NOV-1995 (Rel. 32, Created)				
DT	01-NOV-1995 (Rel. 32, Last sequence update)				
DT	15-JUL-1999 (Rel. 38, Last annotation update)				
DE	BIFUNCTIONAL DIHYDROFOLATE REDUCTASE-THYMIDYLATE SYNTHASE 1 (DHFR-TS)				
DE	[INCLUDES: DIHYDROFOLATE REDUCTASE (EC 1.5.1.3); THYMIDYLATE SYNTHASE (EC 2.1.1.45)]				
GN	THY-1.				
OS	Arabidopsis thaliana (Mouse-ear cross).				
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;				
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;				
OC	Brassicales; Brassicaceae; Arabidopsids.				
OX	NCBI_TaxID=3702;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CV. LANDSBERG RECTA;				
RX	MEDLINE=93386189; PubMed=8374616;				
RA	Lazar G., Zhang H., Goodman H.M.;				
RT	"The origin of the bifunctional dihydrofolate reductase-thymidylate synthase isoenzymes of Arabidopsis thaliana.";				
RL	Plant J. 3:657-668(1993).				
CC	- FUNCTION: TS IS EXCLUSIVELY INVOLVED IN DE NOVO DTMP BIOSYNTHESIS.				
CC	DHFR CAN HAVE TWO DIFFERENT ROLES DEPENDING ON THE SOURCE OF DIHYDROFOLATE: DE NOVO SYNTHESIS OF TETRAHYDROFOLATE OR RECYCLING OF THE DIHYDROFOLATE RELEASED AS ONE OF THE END PRODUCTS OF THE TS CATALYZED REACTION.				
CC	- CATALYTIC ACTIVITY: 5,6,7,8-TETRAHYDROFOLATE + NADP(+) =				
CC	7,8-DIHYDROFOLATE + NADPH.				
CC	- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + DUMP =				
CC	DIHYDROFOLATE + DTMP.				
CC	- PATHWAY: ESSENTIAL STEP FOR DE NOVO GLYCINE AND PURINE SYNTHESIS, DNA PRECURSOR SYNTHESIS, AND FOR THE CONVERSION OF DUMP TO DTMP.				
CC	- SUBUNIT: HETERODIMER OR HOMODIMER (BY SIMILARITY).				
CC	- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE DIHYDROFOLATE REDUCTASE FAMILY.				
CC	- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE THYMIDYLATE SYNTHASE FAMILY.				

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: L08593; AAA32788.1; --
HSSP: P00470; IAN5.
InterPro: IPR000398; --
InterPro: IPR001796; --
Pfam: PF00186; DiHfolate_red; 1.
Pfam: PF00303; thymidylat_synt; 1.
PRINTS: PR00070; DHFR.

DR PRINTS: PR00108; THYMSINTHASE.
 DR PROSITE: PS00075; DHER: 1.
 DR PROSITE: PS00091; THYMIDYLATE SYNTHASE; 1.
 KW Multifunctional enzyme; Oxidoreductase; Transferase; NADP;
 KW Methyltransferase; Nucleotide biosynthesis; One-carbon metabolism;
 KW Multigene family.
 FT DOMAIN 1 200 DIHYDROFOLATE REDUCTASE.
 FT DOMAIN 201 234 HINGE.
 FT DOMAIN 235 519 THYMIDYLATE SYNTHASE.
 FT ACT_SITE 401 401 BY SIMILARITY.
 SQ SEQUENCE 519 AA; 58194 MW; 3A55B8D18998F5AA CRC64;

Query Match 58.8%; Score 47; DB 1; Length 519;
 Best Local Similarity 46.7%; Pred. No. 7.4;
 Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 QLIDIVDLKKNVND 15
 |||:|:|:|:|:|:|
 DB 366 QLVDVLDKKNPDD 380

RESULT 2
 PGM2_YEAST
 ID PGM2_YEAST STANDARD; PRT; 569 AA.
 AC P37012.
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PS2HOGUUCOMUTASE 2 (EC 5.4.2.2) (GLUCOSE PHOSPHOMUTASE 2) (PGM 2).
 GN PGM2 OR GAL5 OR YMR105C OR YMR718.04C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94164176; PubMed=8119301;
 RA Boles E., Liebetrau W., Hofmann M., Zimmermann F.K.;
 RT "A family of hexosephosphate mutases in Saccharomyces cerevisiae.";
 RL Eur. J. Biochem. 220:83-96(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Fu L., Bounelis P., Dey N., Browne B.L., Marchase R.B., Bedwell D.M.;
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RC Hunt S., Bowman S., Barrell B.G., Rajandream M.A.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: THIS ENZYME PARTICIPATES IN BOTH THE BREAKDOWN AND
 CC SYNTHESIS OF GLUCOSE. CAN ALSO ACT ON MANNULOSE.
 CC -!- CATALYTIC ACTIVITY: ALPHA-D-GLUCOSE 1-PHOSPHATE = ALPHA-D-GLUCOSE
 CC 6-PHOSPHATE.
 CC -!- COFACTOR: REQUIRES MAGNESIUM FOR ACTIVITY (BY SIMILARITY).
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -!- SIMILARITY: TO OTHER PHOSPHOGLUCOMUTASES AND PHOSPHOMANNOMUTASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X74823; CAA52820.1; -
 CC EMBL; U09499; AAA91282.1; -
 CC EMBL; Z49702; CAA89741.1; -
 CC PIR; S40264; S40264.
 CC PIR; S41200; S41200.
 CC HSSP; P00949; 1LXT.

DR SGD: S0004711; PGM2.
 DR InterPro: IPR001485; -
 DR Pfam: PF00408; PGM_PMM; 1.
 DR PRINTS: PR00509; PGM_PMM.
 DR PROSITE: PS00710; PGM_PMM; 1.
 KW Isomerase; Phosphorylation; Magnesium; Multigene family.
 FT ACT_SITE 119 119 FORMS THE PHOSPHOSERINE INTERMEDIATE
 FT (BY SIMILARITY).
 SQ SEQUENCE 569 AA; 63088 MW; 45B78AFF8197645E CRC64;

Query Match 57.5%; Score 46; DB 1; Length 569;
 Best Local Similarity 60.0%; Pred. No. 12;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 LIDIVDLKKNVNDL 16
 |||:|:|:|:|:|
 DB 183 LVDIITDKYVNF 197

RESULT 3
 TRPC_LACLA
 ID TRPC_LACLA STANDARD; PRT; 264 AA.
 AC Q01999;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE INDOLE-3-GLYCEROL PHOSPHATE SYNTHASE (BC 4.1.1.48) (IGPS).
 GN TRPC.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TL1403;
 RX MEDLINE=93015708; PubMed=1400208;
 RA Bardowski J., Ehrlich S.D., Chopin A.;
 RT "Tryptophan biosynthesis genes in Lactococcus lactis subsp. lactis.";
 RL J. Bacteriol. 174:6563-6570(1992).
 CC -!- CATALYTIC ACTIVITY: 1-(2-CARBOXYPHENYLAMINO)-1-DEOXY-D-RIBULOSE
 CC 5-PHOSPHATE = 1-(INDOL-3-YL)GLYCEROL 3-PHOSPHATE + CO(2) + H(2)O.
 CC -!- PATHWAY: FOURTH STEP IN BIOSYNTHESIS OF TRYPTOPHAN.
 CC -!- SIMILARITY: BELONGS TO THE TRPC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M87483; AAA25226.1; -
 CC PIR; S35127; S35127.
 CC HSSP; P00909; 1PII.
 CC InterPro: IPR001468; -
 CC Pfam: PF00218; IGPS; 1.
 CC PROSITE: PS00614; IGPS; 1.
 KW Tryptophan biosynthesis; Lyase; Decarboxylase.
 SQ SEQUENCE 264 AA; 29732 MW; E7096EE9FA662535 CRC64;

Query Match 56.2%; Score 45; DB 1; Length 264;
 Best Local Similarity 80.0%; Pred. No. 7.3;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 IDIVDLKKNY 12
 |||:|:|:|:|:|
 DB 68 VDIVDQAKNY 77

RESULT 4

Baker R.T., Varshavsky A.;
"Yeast N-terminal amidase. A new enzyme and component of the N-end
rule pathway.";
J. Biol. Chem. 270:12065-12074 (1995).
CC -I- SIMILARITY: TO YEAST MNNA.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC	EMBL: Z49561; CAA89589.1; -	E -> K (IN REF. 2).
DR	EMBL: L47993; AAB39287.1; -	N -> Y (IN REF. 2).
DR	EMBL: L35584; AAB59321.1; -	SEQUENCE 935 AA; 108427 MW; 4265763FCB4CC6EE CRC64;
DR	PIR: S47939; S47939.	
DR	Hypothetical protein.	
KW	CONFLICT 197 197	
FT	CONFLICT 312 312	
FT	CONFLICT 312 312	
SQ	SEQUENCE 935 AA; 108427 MW; 4265763FCB4CC6EE CRC64;	

```

Query Match      55.0%;   Score 44;   DB 1;   Length 935;
Best Local Similarity 53.3%;   Pred. No. 38;
Matches . 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY    1 QLIDIVDQLKNVND 15
Db    118 QVLDSVDLRYND 132

```

RESULT	7
YJ40_YEAST	
ID	YJ40_YEAST
AC	P47120;
DT	01-FEB-1996 (Rel. 33, Created)
DT	01-FEB-1996 (Rel. 33, Last sequence update)
DT	01-OCT-1996 (Rel. 34, Last annotation update)
DE	HYPOTHETICAL 36.2 KDA PROTEIN IN HAM1-PEM2 INTERGENIC REGION.
GN	YJR070C OR J1814.
OS	Saccharomyces cerevisiae (Baker's yeast).
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;	[1]
RN	SEQUENCE FROM N.A.
RP	

RC STRAIN=3288C;
RX MEDLINE=96437976; PubMed=8840504;
RA Huang M.-E.; Manus V.; Chuat J.-C.; Gallibert F.;
RT "Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open
RT reading frames and a gene cluster with a counterpart on chromosome
RL XI.";
RL Yeast 12:869-875(1996).
RL

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@sib-sib.ch).

CC	-----
DR	EMBL; Z49570; CAA89598.1; -;
DR	EMBL; L47993; AAB39296.1; -;
DR	SGD; S0003831; YJR070C.
KW	Hypothetical protein.
SQ	SEQUENCE 325 AA; 36164 MW; AB36A73B40466CD5 CRC64;

Query Match 53.8%; Score 43; DB 1; Length 325;
Best Local Similarity 58.3%; Pred. No. 18;

	Matches	7;	Conservative	4;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	4	DIVDQLKNYND	15							
		::::	::::							
Db	283	EWVDVLKSYLND	294							

RESULT	8
DRTS_MAIZE	
ID	DRTS_MAIZE
STANDARD;	PRT; 521 AA.
AC	O81395;
DT	30-MAY-2000 (Rel. 39, Created)
DT	30-MAY-2000 (Rel. 39, Last sequence update)
DT	30-MAY-2000 (Rel. 39, Last annotation update)
DE	BIFUNCTIONAL DIHYDROFOLATE REDUCTASE-THYMIDYLATE SYNTHASE (DHER-7S)
DE	[INCLUDES: DIHYDROFOLATE REDUCTASE (EC 1.5.1.3); THYMIDYLATE SYNTHASE (EC 2.1.1.45)].
GN	DRTS.
OS	zea mays (Maize).
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC	Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC	Andropogoneae; Zea.
OX	NCBI_TaxID=4577;

KN	SEQUENCE FROM N.A.
RP	Cox K.M., Robertson D., Fites R.C.;
RA	"Cloning of a cDNA encoding the bifunctional dihydrofolate reductase-
RT	thymidylate synthase (DHFR-TS) in sea mays.;"
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBD databases.
CC	-1- CATALYTIC ACTIVITY: 5,6,7,8-TETRAHYDROFOLATE + NADP(+) =
CC	7,8-DIHYDROFOLATE + NADPH.
CC	-1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + DUMP =
CC	DIHYDROFOLATE + DUMP.
CC	-1- PATHWAY: ESSENTIAL STEP FOR DE NOVO GLYCINE AND PURINE SYNTHESIS,
CC	DNA PRECURSOR SYNTHESIS, AND FOR THE CONVERSION OF DUMP TO DTMP.
CC	-1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC	DIHYDROFOLATE REDUCTASE FAMILY.
CC	-1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE THYMIDYLATE
CC	SYNTASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC	EMBL: AF073488; AAC26003.1; --	
DR	Mendel; 32032; Zeama; 2610; 32032.	
DR	InterPro: IPR000398; --	
DR	InterPro: IPR001796; --	
DR	Pfam; PF00186; Dihfolate_red; 1.	
DR	Pfam; PF00303; thymidylat_synt; 1.	
DR	PRINTS: PR00070; DHFR	
DR	PRINTS: PR00108; THYMSNTHASE.	
DR	PROSITE; PS00075; DHFR; 1.	
DR	PROSITE; PS00091; THYMYDLATE_SYNTHASE; 1.	
KW	Multifunctional enzyme; Oxidoreductase; Transferase; NADP;	
KW	Methyltransferase; Nucleotide biosynthesis; One-carbon metabolism.	
FT	DOMAIN 1 196	DIHYDROFOLATE REDUCTASE.
FT	DOMAIN 197 521	THYMYDLATE SYNTHASE.
FT	ACT_SITE 403 403	BY SIMILARITY.
SQ	SEQUENCE 521 AA; 53965 MW; 81266F8652625F06 CRC64;	

Query Match	53.8%	Score 43:	DB 1:	Length 521:
Best Local Similarity	46.7%	Pred. No. 30:		
Matches 7: Conservative	5:	Mismatches	3:	Indels 0:
Matches 7: Conservative	5:	Mismatches	3:	Indels 0:
Matches 7: Conservative	5:	Mismatches	3:	Indels 0:

QY 1 QLIDIVDLKYNVD 15
||:|:|:|:|:|
Db 368 QLMDVIDIKNDPED 382

```
RESULT 9
ID TSY_BPT4 STANDARD; PRT; 286 AA.
AC P00471;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE THYMIDYLATE SYNTHASE (EC 2.1.1.45) (TS).
GN TD.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae;
OC T4-like phages.
OX NCBI_TaxID=10665;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=84221902; PubMed=6328492;
RA Chu F.K., Maley G.F., Maley F., Belfort M.;
RT "Intervening sequence in the thymidylate synthase gene of
bacteriophage T4.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:3049-3053(1984).
RN [2]
SEQUENCE OF 1-42 FROM N.A.
RX MEDLINE=84212439; PubMed=6327673;
RA Purohit S., Mathews C.K.;
RT "Nucleotide sequence reveals overlap between T4 phage genes encoding
dihydrofolate reductase and thymidylate synthase.";
RL J. Biol. Chem. 259:6261-6266(1984).
RN [3]
SEQUENCE OF 179-188 FROM N.A.
RX MEDLINE=86189927; PubMed=3698096;
RA Chu F.K., Maley G.F., West D.K., Belfort M., Maley F.;
RT "Characterization of the intron in the phage T4 thymidylate synthase
gene and evidence for its self-excision from the primary
transcript.";
RL Cell 45:157-166(1986).
RN [4]
SEQUENCE OF 251-286 FROM N.A.
RX MEDLINE=89034091; PubMed=2846540;
RA Tseng M.J., Hilfinger J.M., Walsh A., Greenberg G.R.;
RT "Total sequence, flanking regions, and transcripts of bacteriophage
T4 nrdA gene, coding for alpha chain of ribonucleoside diphosphate
reductase.";
RL J. Biol. Chem. 263:16242-16251(1988).
RN [5]
X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS).
RX MEDLINE=95101640; PubMed=7803410;
RA Finer-Moore J.S., Maley G.F., Maley F., Montfort W.R., Stroud R.M.;
RT "Crystal structure of thymidylate synthase from T4 phage: component
of a deoxynucleoside triphosphate-synthesizing complex.";
RL Biochemistry 33:15459-15468(1994).
CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + DUMP =
DIHYDROFOLATE + DTMP.
CC -1- PATHWAY: DEOXYRIBONUCLEOTIDE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER.
CC -1- MISCELLANEOUS: THIS ENZYME IS ALSO EXPRESSED BY THE THYVA GENE OF
E. COLI. THE PHAGE AND HOST SYNTHASES EXHIBIT STRIKING
DISSIMILARITIES IN BOTH STRUCTURE AND FUNCTION.
CC -1- SIMILARITY: BELONGS TO THE THYMIDYLATE SYNTHASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; K02035; AAA32492.1;
CC EMBL; K01804; AAA32492.1; JOINED.
CC EMBL; M12742; AAC12816.1;
CC EMBL; J03968; AAA32525.1;
-----
```

```
DR PIR; A00550; SYBPT4.
DR PDB; ITIS; 30-APR-94.
DR InterPro; IP000398;
DR Pfam; PF00303; thymidylat_synt; 1.
DR PRINTS; PRO0108; THYMIDYLATE_SYNTHASE.
DR PROSITE; PS00091; THYMIDYLATE_SYNTHASE; 1.
KW Transferase; Methyltransferase; Nucleotide biosynthesis; 3D-structure.
FT ACT_SITE 136
FT HELIX 3
FT TURN 12
FT TURN 13
FT STRAND 14
FT STRAND 17
FT STRAND 18
FT STRAND 26
FT STRAND 30
FT STRAND 33
FT STRAND 37
FT HELIX 38
FT HELIX 40
FT HELIX 53
FT TURN 55
FT TURN 65
FT STRAND 68
FT STRAND 69
FT HELIX 77
FT TURN 86
FT HELIX 88
FT TURN 92
FT TURN 95
FT HELIX 96
FT TURN 98
FT TURN 99
FT TURN 102
FT STRAND 104
FT STRAND 114
FT STRAND 116
FT STRAND 117
FT STRAND 118
FT STRAND 119
FT HELIX 121
FT TURN 121
FT STRAND 133
FT STRAND 139
FT STRAND 150
FT HELIX 157
FT STRAND 164
FT STRAND 168
FT TURN 179
FT TURN 180
FT HELIX 184
FT TURN 203
FT STRAND 205
FT TURN 224
FT TURN 227
FT STRAND 241
FT HELIX 247
FT HELIX 254
FT TURN 263
FT HELIX 266
FT STRAND 269
FT STRAND 270
SQ SEQUENCE 286 AA; 33073 MW; 9F60863145F5219D CRC64;

Query Match 52 5%; Score 42; DB 1; Length 286;
Best Local Similarity 40.0%; Pred. No. 22;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 QLIDIVDQLKNYVD 15
DB 121 QIIEVIDRIKLPND 135
I::::::
II

RESULT 10
TSYV_ASCSU
ID TSYV_ASCSU STANDARD; PRT; 294 AA.
AC O96650;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE THYMIDYLATE SYNTHASE (EC 2.1.1.45) (TS).
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE FROM N.A.
```


RA Tian L., Tam J.W.O.:
RT "Molecular cloning of thymidylate synthase from Ascaris suum.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + DUMP =
CC DIHYDROFOLATE + DTMP.
CC
CC -!- SUBUNIT: DEOXYRIBONUCLEOTIDE BIOSYNTHESIS.
CC
CC -!- SIMILARITY: BELONGS TO THE THYMIDYLATE SYNTHASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF099672; AAC97507.1; -
DR InterPro: IPR000398; -
DR Pfam: PF00303; thymidylat_synt; 1.
DR PRINTS: PR00108; THYMSNTSHASE.
DR PROSITE: PS00091; THYMIDYLATE_SYNTHASE; 1.
KW Transferase; Methyltransferase; Nucleotide biosynthesis.
FT ACT SITE 176 176 BY SIMILARITY.
SQ SEQUENCE 294 AA; 33053 MW; CDFD4556972A3C50 CRC64;

Query Match 52.5%; Score 42; DB 1; Length 294;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QLIDIVDLKKNV 14
DB 141 QLADVIEQIKNPN 154

RESULT 11
HMD_METKA ID HMD_METKA STANDARD; PRT; 358 AA.
AC Q02394;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE H(2)-FORMING N5,N10-METHYLENETETRAHYDROMETHANOPTERIN DEHYDROGENASE
DE (EC 1.12.99.-) (H2-DEPENDENT METHYLENE-HAMPT DEHYDROGENASE).
GN HMD.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyrales; Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92394151; PubMed=1521540;
RA Ziringibi C., van Dongen W., Schworer B., von Buenau R.,
RA Richter M., Klein A., Thauer R.K.;
RT "H2-forming methylenetetrahydromethanopterin dehydrogenase, a novel
RT type of hydrogenase without iron-sulfur clusters in methanogenic
RT archaea.";
RL Eur. J. Biochem. 208:511-520(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97296404; PubMed=9151968;
RA Klein A.R., Thauer R.K.;
RT "Overexpression of the coenzyme-F420-dependent N5,N10-
RT methylenetetrahydromethanopterin dehydrogenase gene from the
RT hyperthermophilic Methanopyrus kandleri.";
RL Eur. J. Biochem. 245:386-391(1997).
RN [3]
RP CHARACTERIZATION, AND SEQUENCE OF N-TERMINUS.
RX MEDLINE=92125912; PubMed=1772345;
RA Ma K., Ziringibi C., Linder D., Stetter K.O., Thauer R.K.;
RT "N5, N10-methylenetetrahydromethanopterin dehydrogenase (H2-forming)
RT from the extreme thermophile Methanopyrus kandleri.";
RL Arch. Microbiol. 156:43-48(1991).

CC -!- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROMETHANOPTERIN +
CC H(+) = 5,10-METHYLENETETRAHYDROMETHANOPTERIN + H(2).
CC -!- COFACTOR: ZINC.
CC -!- PATHWAY: INVOLVED IN METHANOGENESIS.
CC -!- SUBUNIT: HOMODIMER.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X60719; CAA43127.1; -
DR EMBL: Y10251; CAA71299.1; -
DR PIR: S30603; S30603.
KW Oxidoreductase; Methanogenesis; Zinc.
SQ SEQUENCE 358 AA; 39038 MW; 9B8491B2A5A20771 CRC64;

Query Match 52.5%; Score 42; DB 1; Length 358;
Best Local Similarity 46.2%; Pred. No. 28;
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 4 DIVDQLKNVNDL 16
DB 102 DVMPKLRDYNVDI 114
::: :::::

RESULT 12
PRIM_HELPY ID PRIM_HELPY STANDARD; PRT; 559 AA.
AC P56064;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE DNA PRIMASE (EC 2.7.7.-).
GN DNAG OR HP0012.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kervatage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
RN [2]
RP FUNCTION: DNA PRIMASE IS THE POLYMERASE THAT SYNTHESIZES SMALL
RX RNA PRIMERS FOR THE OKAZAKI FRAGMENTS ON BOTH TEMPLATE STRANDS AT
CC REPLICATION FORKS DURING CHROMOSOMAL DNA SYNTHESIS.
CC
CC -!- COFACTOR: BINDS ONE ZINC ION PER MOLECULE (BY SIMILARITY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL: AE00523; AAD07082.1; -
 DR TIGR: HP0012; -
 DR InterPro: IPR002694; -
 DR InterPro: IPR002936; -
 DR Pfam: PF01751; Toprim; 1.
 DR Pfam: PF01807; zf-CHC2; 1.
 DR KW Transferase; DNA replication; DNA-directed RNA polymerase; Primosome;
 KW Zinc; Metal-binding. 61
 FT ZN_FING 37 CHC2-TYPE (BY SIMILARITY).
 SQ SEQUENCE 559 AA; 63723 MW; E5EB7F2F0BB38170 CRC64;
 Query Match 52.5%; Score 42; DB 1; Length 559;
 Best Local Similarity 57.1%; Pred. No. 45;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 QLIDIVDQKNYVN 14
 Db 12 QTIDIVEISSYVN 25
 RESULT 13
 ID TOP2_PLAFK STANDARD; PRT; 1398 AA.
 AC P41001;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE DNA TOPOISOMERASE II (EC 5.99.1.3).
 GN TOP2.
 OS Plasmodium falciparum (isolate K1 / Thailand).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5839;
 RN [1]
 RP MEDLINE=94316496; PubMed=8041616;
 RA Cheesman S., McAleese S., Goman M., Johnson D., Horrocks P.,
 RA "Ridley R.G., Kilbey B.J.;
 RT "The gene encoding topoisomerase II from Plasmodium falciparum.";
 RL Nucleic Acids Res. 22:2547-2551(1994).
 CC -1- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
 CC BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
 CC MAKES DOUBLE-STRAND BREAKS.
 CC -1- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING
 CC OF DOUBLE-STRANDED DNA.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
 CC NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
 CC RELAX ONLY NEGATIVE SUPERCOILS.
 CC -1- SIMILARITY: BELONGS TO TYPE II TOPOISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X79345; -; NOT_ANNOTATED_CDS.
 DR HSSP: P06786; 1BGW.
 DR InterPro: IPR000947; -
 DR InterPro: IPR001154; -
 DR InterPro: IPR001241; -
 DR InterPro: IPR002205; -
 DR Pfam: PF00204; DNA_topoisoi; 1.
 DR Pfam: PF00521; DNA_topoisoi; 1.
 DR PRINTS; PR00418; TP12FAMILY.
 DR PRINTS; PR00615; CCAATSUBUNTA.
 DR PRINTS; PR01158; TOPTISMRASEII.
 DR PROSITE; PS00177; TOPOISOMERASE-II; 1.
 DR KW Isomerase; Topoisomerase; DNA-binding; ATP-binding; Nuclear protein.

FT NP_BIND 144 149 ATP (POTENTIAL).
 FT ACT_SITE 830 830 DNA CLEAVAGE (BY SIMILARITY).
 FT DOMAIN 271 281 POLY-ASN.
 FT DOMAIN 308 316 POLY-ASN.
 FT DOMAIN 1089 1093 POLY-LYS.
 FT DOMAIN 1227 1234 POLY-LYS.
 SQ SEQUENCE 1398 AA; 161029 MW; BAAD7BEE88FE5BE9 CRC64;
 Query Match 52.5%; Score 42; DB 1; Length 1398;
 Best Local Similarity 54.5%; Pred. No. 1.2e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 Qy 4 DIVDQKNYVN 14
 Db 896 DIIDNIKRYIN 906
 RESULT 14
 ID RP5M_THIFE STANDARD; PRT; 78 AA.
 AC P24694;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE PROBABLE SIGMA(54) MODULATION PROTEIN (ORF3) (FRAGMENT).
 OS Thiobacillus ferrooxidans.
 OC Bacteria; Proteobacteria; gamma subdivision; Acidithiobacillus.
 OX NCBI_TaxID=920;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33020;
 RX MEDLINE=90330545; PubMed=2198257;
 RA Berger D.K., Woods D.R., Rawlings D.E.;
 RT "Complementation of Escherichia coli sigma 54 (NtrA)-dependent
 RT formate hydrogenlyase activity by a cloned Thiobacillus ferrooxidans
 RT ntrA gene.";
 RL J. Bacteriol. 172:4399-4406(1990).
 CC -1- FUNCTION: PROBABLY MODULATES THE ACTIVITY OF SIGMA(54) (RPN).
 CC -1- SIMILARITY: BELONGS TO THE SIGMA(54) MODULATION PROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M58480; AAA27380.1; -
 DR PIR: C37761; C37761.78
 FT NON_TER 78
 SQ SEQUENCE 78 AA; 8751 MW; F39C7C4944E413EB CRC64;
 Query Match 51.2%; Score 41; DB 1; Length 78;
 Best Local Similarity 40.0%; Pred. No. 8.2;
 Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 QLIDIVDQKNYVN 15
 Db 8 QHLDLTDISKYVADE 22
 RESULT 15
 ID DRT2_ARATH STANDARD; PRT; 518 AA.
 AC Q05763;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE BIFUNCTIONAL DIHYDROFOLATE REDUCTASE-THYMIDYLATE SYNTHASE 2 (DHFR-TS)
 DE [INCLUDES: DIHYDROFOLATE REDUCTASE (EC 1.5.1.3); THYMIDYLATE SYNTHASE

(EC 2.1.1.45)].
THV-2 OR AT4G34570 OR T4L20.150.
Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LANDSBERG ERRECTA;
RX MEDLINE=93386189; PubMed=8374616;
RA Lazar G., Zhang H., Goodman H.M.;
RT "The origin of the bifunctional dihydrofolate reductase-thymidylate
RL Plant J. 3:657-668(1993)."
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansorge W., Brandt P., Grivell L., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hobeisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weitzenegger T., Bothe G., Ransperger U., Hilbert H., Braun M.,
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Moolman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,
RA Pettett A., Rajandream M.-A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loeinert T.-H.,
RA Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA Neumann S., Argirou A., Vitale D., Liquori R., Piravandi E.,
RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lechary A., Aubourg S.,
RA Chedford F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacón D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bietke C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Senkon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Spith J., Ryan E., Andrews S., Geisel C., Hillier L.,
RA Nelson J., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshua C.,
RA Antonoli B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Marra M., Martienssen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
thaliana.";
RL Nature 402:769-777(1999).
CC -!- FUNCTION: TS IS EXCLUSIVELY INVOLVED IN DE NOVO DTPM BIOSYNTHESIS.
CC DHPH CAN HAVE TWO DIFFERENT ROLES DEPENDING ON THE SOURCE OF
CC DIHYDROFOLATE: DE NOVO SYNTHESIS OF TETRAHYDROFOLATE OR RECYCLING
CC OF THE DIHYDROFOLATE RELEASED AS ONE OF THE END PRODUCTS OF THE TS
CC CATALYZED REACTION.
CC -!- CATALYTIC ACTIVITY: 5,6,7,8-TETRAHYDROFOLATE + NADP(+) =
CC 7,8-DIHYDROFOLATE + NADPH.
CC -!- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + DUMP =
CC

CC DIHYDROFOLATE + DTPM.
CC -!- PATHWAY: ESSENTIAL STEP FOR DE NOVO GLYCINE AND PURINE SYNTHESIS,
CC DNA PRECURSOR SYNTHESIS, AND FOR THE CONVERSION OF DUMP TO DTPM.
CC -!- SUBUNIT: HETERODIMER OR HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC DIHYDROFOLATE REDUCTASE FAMILY
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE THYMIDYLATE
CC SYNTHASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L08594; AAA32789.1; -;
CC EMBL: AL023094; CAA18836.1; -;
CC EMBL: AL161585; CAB80174.1; -;
CC HSSP: P00470; IANS.
CC InterPro: IPR000398; -;
CC InterPro: IPR001796; -;
CC Pfam: PF00186; Dihfolate_red; 1.
CC Pfam: PF00303; thymidylat_synt; 1.
CC PRINTS: PR00070; DHPH.
CC PRINTS: PR00108; THYMSNTHASE.
CC PROSITE: PS00075; DHPH; 1.
CC PROSITE: PS00091; THYMIDYLATE SYNTHASE; 1.
CC Multifunctional enzyme; Oxidoreductase; Transferase; NADP;
CC Methyltransferase; Nucleotide biosynthesis; One-carbon metabolism;
CC Multigene family.
CC DOMAIN 1 197 DIHYDROFOLATE REDUCTASE.
CC DOMAIN 198 233 HINGE.
CC DOMAIN 234 518 THYMIDYLATE SYNTHASE.
CC ACT_SITE 400 400 BY SIMILARITY.
CC SEQUENCE 518 AA; 57939 MW; A727D38C56CBDB1 CRC64;

Query Match 51.2%; Score 41; DB 1; Length 518;
Best Local Similarity 40.0%; Pred. No. 59;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 OLIDIVDQLKNYND 15
Db 365 QLDDVINKKNPPD 379
||:|||||:|
- - - - -
Search completed: May 23, 2001, 11:20:25
Job time: 523 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:19:43 ; Search time 118.04 Seconds
(without alignments)
15.887 Million cell updates/sec

Title: US-09-522-217-2_COPY_41_56
Perfect score: 80
Sequence: 1 QLIDIVDQKNVNDL 16

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL15:
1: sp-archaea:
2: sp-bacteria:
3: sp-fungi:
4: sp-human:
5: sp-invertebrate:
6: sp-mammal:
7: sp-mhc:
8: sp-organelle:
9: sp-phage:
10: sp-plant:
11: sp-rodent:
12: sp-unclassified:
13: sp-vertebrate:
14: sp-virus:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	58.8	243	2 Q51582	Q51582 plectonema
2	47	58.8	519	10 Q9SIW4	Q9SIW4 arabidopsis
3	45	56.2	549	5 Q9XZV8	Q9XZV8 hydra atten
4	44	55.0	213	5 Q23520	Q23520 caenorhabdi
5	44	55.0	279	14 Q9J869	Q9J869 spodoptera
6	44	55.0	545	3 Q9Y873	Q9Y873 candida alb
7	44	55.0	999	5 Q9U5A3	Q9U5A3 plasmodium
8	43	53.8	189	2 Q9PI54	Q9PI54 campylobact
9	43	53.8	391	5 Q9VVQ7	Q9VVQ7 drosophila
10	43	53.8	608	14 Q9YVW2	Q9YVW2 melanoplus
11	43	53.8	771	10 Q82334	Q82334 arabidopsis
12	43	53.8	782	10 Q9LKR5	Q9LKR5 arabidopsis
13	42	52.5	235	5 P91252	P91252 caenorhabdi
14	42	52.5	957	5 Q9W2N6	Q9W2N6 drosophila
15	42	52.5	2340	5 Q97298	Q97298 plasmodium
16	41	51.2	104	5 Q9V5V1	Q9V5V1 drosophila
17	41	51.2	183	5 Q21131	Q21131 caenorhabdi
18	41	51.2	243	1 Q27016	Q27016 methanobact
19	41	51.2	412	2 Q52055	Q52055 enterococcu

20	41	51.2	462	2 P72665	P72665 synecocyst
21	41	51.2	487	4 Q9NQ10	Q9NQ10 homo sapien
22	41	51.2	498	2 Q9KX54	Q9KX54 pseudomonas
23	41	51.2	552	4 Q14776	Q14776 homo sapien
24	41	51.2	573	14 Q69067	Q69067 human herpe
25	41	51.2	649	5 Q20442	Q20442 caenorhabdi
26	41	51.2	688	10 Q9SYD7	Q9SYD7 arabidopsis
27	41	51.2	769	5 Q20440	Q20440 caenorhabdi
28	41	51.2	809	2 Q9JQB5	Q9JQB5 chlamydia p
29	41	51.2	830	14 Q9WT25	Q9WT25 human herpe
30	41	51.2	830	14 Q9QJ31	Q9QJ31 human herpe
31	41	51.2	1279	3 Q59810	Q59810 schizosacch
32	40.5	50.6	312	10 Q9LJ77	Q9LJ77 arabidopsis
33	40.5	50.6	430	5 Q17702	Q17702 caenorhabdi
34	40	50.0	69	2 Q9PR47	Q9PR47 ureaplasma
35	40	50.0	72	1 Q29171	Q29171 archaeoglob
36	40	50.0	99	14 Q9QSV5	Q9QSV5 human immun
37	40	50.0	205	5 Q15795	Q15795 p strain dd
38	40	50.0	206	5 P91253	P91253 caenorhabdi
39	40	50.0	287	2 Q48615	Q48615 lactococcus
40	40	50.0	356	13 Q73679	Q73679 brachydanio
41	40	50.0	359	2 Q9WXP9	Q9WXP9 thermotoga
42	40	50.0	370	14 Q9YVM7	Q9YVM7 melanoplus
43	40	50.0	373	14 Q9Q8Y3	Q9Q8Y3 rabbit fibr
44	40	50.0	427	5 Q9VLS6	Q9VLS6 drosophila
45	40	50.0	540	4 Q9Y6F8	Q9Y6F8 homo sapien

ALIGNMENTS

RESULT 1

ID Q51582 PRELIMINARY; PRT; 243 AA.
AC Q51582;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-MAY-2000 (Tremblrel. 01, Last sequence update)
DE HYPOTHETICAL 27.3 KDA PROTEIN.
OS Plectonema boryanum.
OC Bacteria; Cyanobacteria; Oscillatoriales; Plectonema.
OX NCBI_TaxID=1184;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM-M101;
RA Okuhara H., Matsumura T., Fujita Y., Hase T.;
RT "Cloning and characterization of two distinct genes encoding ferredoxin- and NADH-dependent glutamate synthases in the cyanobacterium, Plectonema boryanum."
RT Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL: D85230; BAA12740.1; -
DR INTERPRO: IPR002198; -
DR PFAM: PF00106; adh_short; 1.
KW Hypothetical protein.
SQ SEQUENCE 243 AA; 27266 MW; 6241CB696414F272 CRC64;

Query Match 58.8%; Score 47; DB 2; Length 243;
Best Local Similarity 58.3%; Pred. No. 16;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QLIDIVDQKNV 12

Db 217 QLIDVIDQLENH 228

RESULT 2

ID Q9SIW4 PRELIMINARY; PRT; 519 AA.
AC Q9SIW4;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)

DE DIHYDROFOLATE REDUCTASE-THYMIDYLATE SYNTHASE.
GN AT2G16370.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell L.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Talion L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome II of Arabidopsis thaliana.";
RL Nature 402:761-768(1999).
DR EMBL; AC007047; AAD22302.1; -.
DR HSSP; P13100; 1C17.
DR INTERPRO; IPR000398; -.
DR INTERPRO; IPR001796; -.
DR PFAM; PF00186; DiHfolate_red; 1.
DR PFAM; PF00303; thymidylat_synt; 1.
DR PRINTS; PR00070; DHFR.
DR PROSITE; PS00108; THYMDSINTHASE.
DR PROSITE; PS00075; DHFR; 1.
DR PROSITE; PS00091; THYMIDYLATE_SYNTHASE; 1.
SQ SEQUENCE 519 AA; 58143 MW; B5EB36A3A936580F CRC64;

Query Match 58.8%; Score 47; DB 10; Length 519;
Best Local Similarity 46.7%; Pred. No. 35;
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 QLIDIVDLKKNVND 15
||:|:|:|:|:|:|
DB 366 QLVDVDRKKNPDD 380

RESULT 3
Q9XZV8 PRELIMINARY; PRT; 549 AA.
AC Q9XZV8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE LAMIN.
GN LAMIN.
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroidea; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6087;
RN [1]
RP SEQUENCE FROM N.A.
RA Erber A., Riemer D., Hofmeister H., Bovenschulte M., Stick R.,
RA Panopoulou G., Lehrach H., Weber K.;
RT "Characterization of the Hydra lamin and its gene; a molecular
RT phylogeny of metazoan lamins.";
RL J. Mol. Evol. 0:0-0(0).
CC -1- SIMILARITY: TO ALL OTHER INTERMEDIATE FILAMENT PROTEINS.
DR EMBL; AJ005934; CAB43346.1; -.
DR INTERPRO; IPR001322; -.
DR INTERPRO; IPR001664; -.
DR PFAM; PF00038; filament; 2.
DR PFAM; PF00932; IF_tail; 1.
DR PROSITE; PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Heptad repeat pattern.
SQ SEQUENCE 549 AA; 63823 MW; 4270CB031DEBE61D CRC64;

Query Match 56.2%; Score 45; DB 5; Length 549;

DE DIHYDROFOLATE REDUCTASE-THYMIDYLATE SYNTHASE.
GN AT2G16370.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell L.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Talion L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome II of Arabidopsis thaliana.";
RL Nature 402:761-768(1999).
DR EMBL; AC007047; AAD22302.1; -.
DR HSSP; P13100; 1C17.
DR INTERPRO; IPR000398; -.
DR INTERPRO; IPR001796; -.
DR PFAM; PF00186; DiHfolate_red; 1.
DR PFAM; PF00303; thymidylat_synt; 1.
DR PRINTS; PR00070; DHFR.
DR PROSITE; PS00108; THYMDSINTHASE.
DR PROSITE; PS00075; DHFR; 1.
DR PROSITE; PS00091; THYMIDYLATE_SYNTHASE; 1.
SQ SEQUENCE 519 AA; 58143 MW; B5EB36A3A936580F CRC64;

Query Match 58.8%; Score 47; DB 10; Length 519;
Best Local Similarity 46.7%; Pred. No. 35;
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 QLIDIVDLKKNVND 15
||:|:|:|:|:|:|
DB 366 QLVDVDRKKNPDD 380

RESULT 3
Q9XZV8 PRELIMINARY; PRT; 549 AA.
AC Q9XZV8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE LAMIN.
GN LAMIN.
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroidea; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6087;
RN [1]
RP SEQUENCE FROM N.A.
RA Erber A., Riemer D., Hofmeister H., Bovenschulte M., Stick R.,
RA Panopoulou G., Lehrach H., Weber K.;
RT "Characterization of the Hydra lamin and its gene; a molecular
RT phylogeny of metazoan lamins.";
RL J. Mol. Evol. 0:0-0(0).
CC -1- SIMILARITY: TO ALL OTHER INTERMEDIATE FILAMENT PROTEINS.
DR EMBL; AJ005934; CAB43346.1; -.
DR INTERPRO; IPR001322; -.
DR INTERPRO; IPR001664; -.
DR PFAM; PF00038; filament; 2.
DR PFAM; PF00932; IF_tail; 1.
DR PROSITE; PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Heptad repeat pattern.
SQ SEQUENCE 549 AA; 63823 MW; 4270CB031DEBE61D CRC64;

Query Match 56.2%; Score 45; DB 5; Length 549;

Best Local Similarity 56.2%; Pred. No. 74;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 QLIDIVDLKKNVNDL 16
|:|:|:|:|:|:|
DB 155 QVKLEDELKKEYODDL 170

RESULT 4
Q23520 PRELIMINARY; PRT; 213 AA.
AC Q23520;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE SIMILAR TO S-CRYSTALLIN.
GN ZK546.11.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightons J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Hallsworth K.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U29380; AAA68744.1; -.
DR HSSP; P46088; ZGSQ.
DR INTERPRO; IPR000521; -.
DR PFAM; PF00043; GST; 1.
SQ SEQUENCE 213 AA; 24180 MW; 338AC8F76D29FF58 CRC64;

Query Match 55.0%; Score 44; DB 5; Length 213;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 QLIDIVDLKKNVNDL 16
|:|:|:|:|:|:|
DB 87 QVNAIADQFKDYLNDI 102

RESULT 5
Q9J869 PRELIMINARY; PRT; 279 AA.
AC Q9J869;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE ORF66 LEF5.
OS Spodoptera exigua nucleopolyhedrovirus.

OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=10454;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20036646; PubMed=10567663;
RA IJkel W.F.J., van Strien E.A., Heldens J.G., Broer R., Zuidema D.,
RA Goldbach R.W., Vlak J.M.;
RT "Sequence and organization of the spodoptera exigua multicapsid
RT nucleopolyhedrovirus genome.";
RL J. Gen. Virol. 80:3289-3304(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA IJkel W.F.J., van Strien E.A., Heldens J.G.M., Broer R., Zuidema D.,
RA Goldbach R.W., Vlak J.M.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF169823; AAF33596.1;
SQ SEQUENCE 279 AA; 32847 MW; -B0DD4C420B09D264 CRC64;

Query Match 55.0%; Score 44; DB 14; Length 279;
Best Local Similarity 61.5%; Pred. No. 51;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 LIDIVDLKKNYVN 14
| : : : : :
Db 145 LVRRLDQIKNYVN 157

RESULT 6

ID Q9Y873 PRELIMINARY; PRT; 545 AA.
AC Q9Y873;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE ADENYLATE CYCLASE-ASSOCIATED PROTEIN HOMOLOG.
GN HAAL.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Eukaryotic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SC5314;
RA Bahn Y.S., Sundstrom P.;
RT "The role of HAAL gene product in Candida albicans dimorphism.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF163838; AAD42978.1;
DR INTERPRO: IPR001837;
DR PFAM; PF01213; CAP; 1.
DR PROSITE; PS01088; CAP_1; 1.
SQ SEQUENCE 545 AA; 59605 MW; A515AF51A7E5F47E CRC64;

Query Match 55.0%; Score 44; DB 3; Length 545;
Best Local Similarity 46.7%; Pred. No. 1e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 QLIDIVDLKKNYND 15
| : : : : :
Db 233 QTLSEVDELKKNYVKE 247

RESULT 7

ID Q9U5A3 PRELIMINARY; PRT; 999 AA.
AC Q9U5A3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE ERYTHROCYTE BINDING PROTEIN HOMOLOG (FRAGMENT).
OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FCR-3;
RA Alcina A., Matesanz F.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF042160; AAF21647.1;
FT NON_TER 1
FT NON_TER 999
SQ SEQUENCE 999 AA; 118410 MW; F373EDF56DE66B94 CRC64;

Query Match 55.0%; Score 44; DB 5; Length 999;
Best Local Similarity 63.6%; Pred. No. 1.9e+02;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 6 VDOLKKNYNDL 16
| : : : : :
Db 798 VDELKKNYLNLI 808

RESULT 8

ID Q9PI54 PRELIMINARY; PRT; 189 AA.
AC Q9PI54;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PUTATIVE MEMBRANE PROTEIN.
GN CJ0454C.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagsis K., Karlyshev A., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., Van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
DR EMBL; AL139075; CAB75092.1;
SQ SEQUENCE 189 AA; 21834 MW; 04B6D48CC251CAF CRC64;

Query Match 53.8%; Score 43; DB 2; Length 189;
Best Local Similarity 57.1%; Pred. No. 48;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 QLIDIVDLKKNYVN 14
| : : : : :
Db 148 RLYDFIDALKKNYNN 161

RESULT 9

ID Q9VV07 PRELIMINARY; PRT; 391 AA.
AC Q9VV07;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG18231 PROTEIN.
GN CG18231.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brannon R.C., Rogers Y.-H.C., Blazew R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Besson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,

RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laspo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M.C., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

DR EMBL; AE003520; AAF49253.1; -

DR FLYBASE; FBgn0036796; CG18231.

SQ SEQUENCE 391 AA; 45929 MW; 3C73FB16DA29BDAD CRC64;

Query Match 53.8%; Score 43; DB 5; Length 391;

Best Local Similarity 58.3%; Pred. No. 1e+02;

Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 DIVDOLKNVND 15

DB 302 DEIDKLKHLND 313

RESULT 10

Q9YVV2

ID Q9YV2

AC Q9YV2

DT 01-MAY-1999 (Tremblrel. 10, Created)

DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)

DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)

DE ORF MSV140 HYPOTHETICAL PROTEIN.

GN MSV140.

OS Melanoplus sanguinipes entomopoxvirus.

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;

OC Entomopoxvirus B.

OX NCBI_TaxID=83191;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=TUCSON;

RX MEDLINE=99102612; PubMed=9847359;

RA Afonso C.L., Tulman E.R., Lu Z., Oma E., Kutish G.F., Rock D.L.;

RT "The genome of Melanoplus sanguinipes entomopoxvirus.";

RL J. Virol. 73:533-552(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=TUCSON;

RA Afonso C.L., Tulman E.R., Lu Z., Oma E., Kutish G.F., Rock D.L.;

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF063866; AAC97872.1; -

SQ SEQUENCE 608 AA; 73001 MW; 41F78C801CFD0DC9 CRC64;

Query Match 53.8%; Score 43; DB 14; Length 608;

Best Local Similarity 50.0%; Pred. No. 1.6e+02;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 IDIVDOLKNVNDL 16

DB 316 IDIIEIGNYNDI 329

RESULT 11

ID 082334

AC 082334

DT 01-NOV-1998 (Tremblrel. 08, Created)

DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)

DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)

DE HYPOTHETICAL 85.5 KDA PROTEIN.

GN T15J14.1.

OS Arabidopsis thaliana (Mouse-ear cross).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; euroids II;

OC Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,

RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;

RT "Arabidopsis thaliana chromosome II BAC T15J14 genomic sequence.";

RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC005957; AAD03355.1; -

SQ SEQUENCE 771 AA; 85481 MW; 677EIF58096DC455 CRC64;

Query Match 53.8%; Score 43; DB 10; Length 771;

Best Local Similarity 42.9%; Pred. No. 2.1e+02;

Matches 6; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 QLIDIVDOLKNVYN 14

DB 422 ELQEVIDKIKNVN 435

RESULT 12

Q9LKR5

ID Q9LKR5

AC Q9LKR5

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)

DE T26D3.7 PROTEIN.

GN T26D3.7

OS Arabidopsis thaliana (Mouse-ear cross).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; euroids II;

OC Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;

OY 1 QLIDIVDLKKNY 12
|:|:|:|:|:|:|
Db 565 QLIDIFDLKKNY 576

RESULT 15

O97298
ID O97298 PRELIMINARY; PRT: 2340 AA.
AC O97298;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE PFC1015C PROTEIN.
GN PFC1015C.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RA Oliver K., Bowman S., Harris D., Lawson D., Quail M., Barrell B.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL034559; CAB39045.1; -;
SQ SEQUENCE 2340 AA; 280915 MW; 040F69BAF1B8DF10 CRC64;

Query Match 52.5%; Score 42; DB 5; Length 2340;
Best Local Similarity 53.8%; Pred. No. 9.5e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 LIDIVDLKKNY 14
|:|:|:|:|:|:|
Db 1252 LIDFIFQIKKNY 1264

Search completed: May 23, 2001, 11:19:45
Job time: 598 sec

GenCore version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:11:35 ; Search time 108.07 Seconds
(without alignments)
8.463 Million cell updates/sec

Title: US-09-522-217-2_COPY_69_84
Perfect score: 92
Sequence: 1 TNCWSAFSCFQKAQL 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	92	100.0	162 21 B18623	A human zalphall 1
2	92	100.0	519 21 B18627	Amino acid sequenc
3	59	64.1	146 21 B18624	A mouse zalphall 1
4	59	64.1	510 21 B18628	Amino acid sequenc
5	42	45.7	94 22 B55860	PDZ encoded domain
6	42	45.7	94 22 B57651	KIAA0561 protein p
7	42	45.7	94 22 B58061	KIAA0561 protein p
8	41	44.6	42 20 Y14430	Human secreted pro
9	41	44.6	57 19 W75033	Fragment of human
10	41	44.6	81 20 Y35840	Chlamydia pneumoni
11	41	44.6	183 19 W75040	Fragment of human

12	41	44.6	241	21	Y73348	HTFM clone 839651
13	41	44.6	570	18	W36002	Human Fchd531 gene
14	41	44.6	570	21	Y45013	Protein encoded by
15	41	44.6	1111	17	R91428	Kalinin/laminin 5
16	41	44.6	1172	21	B48469	Human laminin 5 po
17	41	44.6	1172	21	B48471	Human laminin 5 po
18	41	44.6	1193	17	R91427	Kalinin/laminin 5
19	41	44.6	1193	21	B48468	Human laminin 5 po
20	41	44.6	1193	21	B48470	Human laminin 5 po
21	41	44.6	4572	19	W52845	A. mediterranei ri
22	40	43.5	363	20	Y37863	Amino acid sequenc
23	39	42.4	26	21	Y81247	Human Chp (Cdc42hs
24	39	42.4	106	19	W77722	Peptide transport
25	39	42.4	129	21	G02608	Human secreted pro
26	39	42.4	148	13	R26410	Human phospholipas
27	39	42.4	148	16	R63054	Human PLA2 type I.
28	39	42.4	156	21	B54059	Human pancreatic c
29	39	42.4	197	21	G09442	Arabidopsis thalia
30	39	42.4	198	21	B25125	Pinus radiata cell
31	39	42.4	205	21	Y81246	Human Chp (Cdc42hs
32	39	42.4	211	21	B25443	Pinus radiata cell
33	39	42.4	236	21	Y81244	Human Chp (Cdc42hs
34	39	42.4	236	21	Y81250	Human Chp (Cdc42hs
35	39	42.4	236	21	Y81251	Human Chp (Cdc42hs
36	39	42.4	384	22	B65609	Novel protein kina
37	39	42.4	699	21	G30475	Arabidopsis thalia
38	39	42.4	708	21	G30474	Arabidopsis thalia
39	39	42.4	752	21	B53398	Human colon cancer
40	39	42.4	836	21	G30473	Arabidopsis thalia
41	39	42.4	996	21	B25514	Pinus radiata cell
42	39	42.4	996	21	B25558	Pinus radiata cell
43	39	42.4	1266	19	W48721	Human cytoplasmic
44	38.5	41.8	368	14	R42448	Lignin peroxidase.
45	38	41.3	22	19	Y20678	Human neurofilamen

ALIGNMENTS

RESULT 1
B18623
ID B18623 standard; Protein; 162 AA.
XX
AC B18623;
XX
XX
DT 22-JAN-2001 (first entry)
XX
XX A human zalphall ligand polypeptide.
DE
XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
XX
XX Homo sapiens.
OS
XX
PN WO200053761-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-US06067.
XX
PR 09-MAR-1999; 99US-0264908.
PR 11-MAR-1999; 99US-0265992.
PR 01-JUL-1999; 99US-0142013.
XX
XX (ZYMO) ZYMOGENETICS INC.
PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Gross JA;
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX WPI; 2000-565600/52.
DR N-PSDB; A75552.
XX
PT New human cytokine, designated zalphall ligand, useful for stimulating

PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX Disclosure; Page 205-206; 256pp; English.
 PS
 CC The present sequence represents a human zalphall ligand polypeptide,
 CC which is a cytokine. The zalphall ligand is useful for stimulating the
 CC proliferation and development of haematopoietic cells in vitro and in
 CC vivo. Zalphall ligand polynucleotides can be used as primers or probes
 CC for cloning the zalphall gene. The zalphall ligand is useful for
 CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
 CC used for treating leukaemias and lymphomas. Antagonists against zalphall
 CC ligand are useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 SQ Sequence 162 AA;

Query Match 100.0%; Score 92; DB 21; Length 162;

Best Local Similarity 100.0%; Pred. No. 6.1e-07;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNCWSAFSCFQKQAL 16

DB 69 tncwsafscfqkql 84

RESULT 2

ID B18627
 AC B18627 standard; Protein; 519 AA.

DE B18627;

DT 22-JAN-2001 (first entry)

DE Amino acid sequence of MBP-human zalphall ligand fusion in pTAP126.

KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.

OS Synthetic.

OS Homo sapiens.

PN WO200053761-A2.

PD 14-SEP-2000.

PF 09-MAR-2000; 2000WO-US06067.

PR 09-MAR-1999; 99US-0264908.

PR 11-MAR-1999; 99US-0265992.

PR 01-JUL-1999; 99US-0142013.

XX (ZYMO) ZYMOGENETICS INC.

XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Gross JA;

PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;

XX WPI: 2000-565600/52.

DR N-PSDB; A75599.

XX New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX

Example 31; Page 233-235; 256pp; English.

XX
 CC The present sequence represents a MBP-human zalphall ligand fusion in
 CC the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
 CC is useful for stimulating the proliferation and development of
 CC haematopoietic cells in vitro and in vivo. Zalphall ligand
 CC polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 SQ Sequence 519 AA;

Query Match 100.0%; Score 92; DB 21; Length 519;

Best Local Similarity 100.0%; Pred. No. 1.9e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNCWSAFSCFQKQAL 16

DB 426 tncwsafscfqkql 441

RESULT 3

ID B18624

AC B18624 standard; Protein; 146 AA.

DE B18624;

DT 22-JAN-2001 (first entry)

DE A mouse zalphall ligand polypeptide.

KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.

OS Mus musculus.

PN WO200053761-A2.

PD 14-SEP-2000.

PF 09-MAR-2000; 2000WO-US06067.

PR 09-MAR-1999; 99US-0264908.

PR 11-MAR-1999; 99US-0265992.

PR 01-JUL-1999; 99US-0142013.

XX (ZYMO) ZYMOGENETICS INC.

XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Gross JA;

PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;

DR WPI: 2000-565600/52.

DR N-PSDB; A75580.

XX New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX

Disclosure; Page 222-223; 256pp; English.

XX The present sequence represents a mouse zalphall ligand polypeptide,
 CC which is a cytokine. The zalphall ligand is useful for stimulating the
 CC proliferation and development of haematopoietic cells in vitro and in

CC vivo. Zalphall ligand polynucleotides can be used as primers or probes
 CC for cloning the zalphall gene. The zalphall ligand is useful for
 CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
 CC used for treating leukaemias and lymphomas. Antagonists against zalphall
 CC ligand are useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.

XX Sequence 146 AA;

Query Match 64.1%; Score 59; DB 21; Length 146;

Best Local Similarity 66.7%; Pred. No. 0.062; Mismatches 4; Indels 0; Gaps 0;

QY 2 NCEWSAFSCFOKAQL 16

Db 63 hcehaafacqkkl 77

RESULT 4

ID B18628 standard; Protein; 510 AA.

AC B18628;

DT 22-JAN-2001 (first entry)

XX Amino acid sequence of MBP-mouse zalphall ligand fusion in pTAP134.

XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;

KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.

XX Synthetic.

OS Mus musculus.

XX WO200053761-A2.

PN 14-SEP-2000.

XX 09-MAR-2000; 2000WO-US06067.

PR 09-MAR-1999; 99US-0264908.

PR 11-MAR-1999; 99US-0265992.

PR 01-JUL-1999; 99US-0142013.

XX (ZYMO) ZYMOGENETICS INC.

PA Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;

PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;

XX WPI: 2000-565600/52.

DR N-PSDB; A75602.

XX New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX Example 31; Page 239-240; 256pp; English.
 XX The present sequence represents a MBP-mouse zalphall ligand fusion in
 CC the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
 CC is useful for stimulating the proliferation and development of
 CC haematopoietic cells in vitro and in vivo. Zalphall ligand
 CC polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating

CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.

XX Sequence 510 AA;

Query Match 64.1%; Score 59; DB 21; Length 510;

Best Local Similarity 66.7%; Pred. No. 0.21; Mismatches 4; Indels 0; Gaps 0;

QY 2 NCEWSAFSCFOKAQL 16

Db 427 hcehaafacqkkl 441

RESULT 5

B55860

ID B55860 standard; Peptide; 94 AA.

AC B55860;

DT 07-MAR-2001 (first entry)

XX PDZ encoded domain #30.

XX Hematopoietic cell; PDZ; PL; autoimmune disease; inflammation;
 KW allergy; asthma; multiple sclerosis; cancer; infection.

XX Synthetic.

XX WO200069896-A2.

PN 23-NOV-2000.

XX 12-MAY-2000; 2000WO-US13161.

XX 14-MAY-1999; 99US-0134114.

PR 14-MAY-1999; 99US-0134117.

PR 21-OCT-1999; 99US-0134118.

PR 29-OCT-1999; 99US-0160860.

PR 13-DEC-1999; 99US-0162498.

PR 14-JAN-2000; 2000US-0176195.

PR 14-FEB-2000; 2000US-0182296.

PR 11-APR-2000; 2000US-0196460.

PR 11-APR-2000; 2000US-0196527.

XX (ARBO-) ARBOR VITA CORP.

XX Lu PS;

XX WPI: 2001-080245/09.

XX Modulating a biological function of an endothelial cell or

PT hematopoietic cell, useful for treating autoimmune diseases and

PT infectious diseases, by administering an antagonist that inhibits

PT binding between a PDZ protein and a PL protein -

XX Disclosure; Page 28-43; 141pp; English.

XX The present invention relates to a new method for modulating a

CC biological function of an endothelial cell or hematopoietic cell. The

CC method involves introducing into a cell, an antagonist that inhibits

CC binding between a PDZ protein and a PL protein. The inhibitor is used

CC to treat a disease mediated by hematopoietic cells, e.g. autoimmune

CC disease. It may also be used to prevent transplantation rejection of
 CC a solid organ transplant. The method may also be used in the treatment
 CC of inflammation, allergy, inflammatory bowel diseases, ulcerative
 CC colitis, ileitis, psoriasis, asthma, atopic dermatitis, autoimmune
 CC diseases (e.g. rheumatoid arthritis, multiple sclerosis,
 CC insulin-dependent diabetes, Hashimoto thyroiditis, osteoarthritis,
 CC graft rejection, transplantation rejection), atherosclerosis, cancers,
 CC infectious diseases, ischemia, vasculitis and Crohn's disease.
 XX
 SQ Sequence 94 AA;

Query Match 45.7%; Score 42; DB 22; Length 94;
 Best Local Similarity 41.7%; Pred. No. 16;
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 3 CEWSAFSCFQKA 14
 | | : : | | : : |
 Db 67 ctwtswscczra 78

RESULT 6
 B57651
 ID B57651 standard; Protein; 94 AA.
 XX
 AC B57651;
 XX
 DT 12-MAR-2001 (first entry)
 XX
 DE KIAA0561 protein PDZ domain.
 XX
 KW Endothelial cell; haematopoietic cell; PDZ domain protein;
 KW PL domain protein; leukocyte activation; synapse formation;
 KW transmembrane neurotransmitter receptor; autoimmune disease;
 KW transplantation rejection; inflammation; allergy;
 KW inflammatory bowel disease; ulcerative colitis; ileitis; psoriasis;
 KW asthma; atopic dermatitis; atherosclerosis; cancer; infectious disease;
 KW ischaemia; vasculitis; Crohn's disease.
 XX
 OS Homo sapiens.
 XX
 PN WO2000069897-A2.
 XX
 PD 23-NOV-2000.
 XX
 PF 12-MAY-2000; 2000WO-US13166.
 XX
 PR 14-MAY-1999; 99US-0134114.
 PR 14-MAY-1999; 99US-0134117.
 PR 14-MAY-1999; 99US-0134118.
 PR 21-OCT-1999; 99US-0160860.
 PR 29-OCT-1999; 99US-0162498.
 PR 13-DEC-1999; 99US-0170453.
 PR 14-JAN-2000; 2000US-0176195.
 PR 14-FEB-2000; 2000US-0182296.
 PR 11-APR-2000; 2000US-0196460.
 PR 11-APR-2000; 2000US-0196527.
 XX
 PA (ARBO-) ARBOR VITA CORP.
 XX
 PI Lu PS;
 XX
 DR WPI; 2001-025003/03.
 XX
 PT New inhibitors of binding of a PDZ protein and PL protein for
 PT inhibiting T cell-mediated response by hematopoietic cells, or for
 PT treating diseases characterized by inflammatory and humoral immune
 PT responses, e.g. inflammation, cancer
 XX
 PS Disclosure; Page 42; 139pp; English.
 XX
 CC The present invention relates to a method for modulating a biological
 CC function of an endothelial cell or haematopoietic cell, comprises

CC introducing into a cell an antagonist that inhibits binding between a
 CC PDZ domain protein and a PL domain protein to result in inhibition of
 CC leukocyte activation. The present sequence is a PDZ domain. PDZ domains
 CC of proteins are named after three prototypical proteins: PSD95,
 CC Drosophila large disc protein and Zonula Occludin 1 protein. PDZ domain
 CC proteins are involved in synapse formation by organising transmembrane
 CC neurotransmitter receptors through intracellular interactions. The
 CC inhibitors identified by the present invention can be used to treat a
 CC disease mediated by haematopoietic cells, e.g. inflammatory bowel disease,
 CC ulcerative colitis, ileitis, psoriasis, respiratory allergic diseases,
 CC (e.g. asthma), atopic dermatitis, autoimmune diseases (e.g. rheumatoid
 CC arthritis, multiple sclerosis, insulin-dependent diabetes, Hashimoto
 CC thyroiditis, osteoarthritis), atherosclerosis, cancers, infectious
 CC diseases (e.g. viral infection), ischaemia, vasculitis and Crohn's
 CC disease. The inhibitors can also be used to prevent transplantation
 CC rejection of a solid organ transplant.
 XX
 SQ Sequence 94 AA;

Query Match 45.7%; Score 42; DB 22; Length 94;
 Best Local Similarity 41.7%; Pred. No. 16;
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 3 CEWSAFSCFQKA 14
 | | : : | | : : |
 Db 67 ctwtswscczra 78

RESULT 7
 B58061
 ID B58061 standard; Protein; 94 AA.
 XX
 AC B58061;
 XX
 DT 12-MAR-2001 (first entry)
 XX
 DE KIAA0561 protein PDZ domain.
 XX
 KW Endothelial cell; haematopoietic cell; PDZ domain protein;
 KW PL domain protein; leukocyte activation; synapse formation;
 KW transmembrane neurotransmitter receptor; autoimmune disease;
 KW transplantation rejection; inflammation; allergy;
 KW inflammatory bowel disease; ulcerative colitis; ileitis; psoriasis;
 KW asthma; atopic dermatitis; atherosclerosis; cancer; infectious disease;
 KW ischaemia; vasculitis; Crohn's disease.
 XX
 OS Homo sapiens.
 XX
 PN WO2000069898-A2.
 XX
 PD 23-NOV-2000.
 XX
 PF 12-MAY-2000; 2000WO-US13205.
 XX
 PR 14-MAY-1999; 99US-0134114.
 PR 14-MAY-1999; 99US-0134117.
 PR 14-MAY-1999; 99US-0134118.
 PR 21-OCT-1999; 99US-0160860.
 PR 29-OCT-1999; 99US-0162498.
 PR 13-DEC-1999; 99US-0170453.
 PR 14-JAN-2000; 2000US-0176195.
 PR 14-FEB-2000; 2000US-0182296.
 PR 11-APR-2000; 2000US-0196460.
 PR 11-APR-2000; 2000US-0196527.
 XX
 PA (ARBO-) ARBOR VITA CORP.
 XX
 PI Lu PS;
 XX
 DR WPI; 2001-061214/07.
 XX

PT Modulating a biological function of a hematopoietic cell for treating
 PT an allergic response, or diseases mediated by immune system cells,
 PT comprises introducing into the cell a PDZ-PL interaction enhancer or
 PT inhibitor -
 XX
 XX
 XX Disclosure; Page 44; 143pp; English.
 PS
 CC The present invention relates to a method for modulating a biological
 CC function of an endothelial cell or hematopoietic cell, comprises
 CC introducing into a cell an antagonist that inhibits binding between a
 CC PDZ domain protein and a PL domain protein to result in inhibition of
 CC leukocyte activation. The present sequence is a PDZ domain. PDZ domains
 CC of proteins are named after three prototypical proteins: PSD95,
 CC Drosophila large disc protein and Zonula Occludin 1 protein. PDZ domain
 CC proteins are involved in synapse formation by organising transmembrane
 CC neurotransmitter receptors through intracellular interactions. The
 CC inhibitors identified by the present invention can be used to treat a
 CC disease mediated by hematopoietic cells, e.g. autoimmune disease,
 CC inflammation, allergy (e.g. drug allergies), inflammatory bowel diseases,
 CC ulcerative colitis, ileitis, psoriasis, respiratory allergic diseases
 CC (e.g. asthma), atopic dermatitis, autoimmune diseases (e.g. rheumatoid
 CC arthritis, multiple sclerosis, insulin-dependent diabetes, Hashimoto
 CC thyroiditis, osteoarthritis), atherosclerosis, cancers, infectious
 CC diseases (e.g. viral infection), ischaemia, vasculitis and Crohn's
 CC disease. The inhibitors can also be used to prevent transplantation
 CC rejection of a solid organ transplant.
 XX
 SQ Sequence 94 AA;

Query Match 45.7%; Score 42; DB 22; Length 94;
 Best Local Similarity 41.7%; Pred. No. 16;
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 3 CEWSAFSCFOKA 14
 Db 67 ctwtswsczra 78
 |:::|::|::|

RESULT 8

Y14430
 ID Y14430 standard; Protein; 42 AA.

XX Y14430;

DT 17-AUG-1999 (first entry)

DE Human secreted protein encoded by gene 20 clone HPEAD79.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX Homo sapiens.

XX WO9919339-A1.

XX 22-APR-1999.

XX 08-OCT-1998; 98WO-US21142.

XX 09-OCT-1997; 97US-0071498.

XX 09-OCT-1997; 97US-0061463.

XX 09-OCT-1997; 97US-0061527.

XX 09-OCT-1997; 97US-0061529.

XX 09-OCT-1997; 97US-0061532.

XX 09-OCT-1997; 97US-0061536.

PA (HUMA-) HUMAN GENOME SCI INC.

XX Brewer LA, Duan R, Ebner R, Ferrie AM, Florence C;

PI Florence KA, Greene JM, Olsen HS, Rosen CA, Ruben SM;

XX Young PE, Yu G;

DR WPI; 1999-277587/23.

DR N-PSDB; X79030.

XX New isolated human genes and the secreted polypeptides they encode

PS Claim 11; Page 192; 226pp; English.

XX This sequence represents a secreted human protein encoded by the gene
 CC clone detailed in the descriptor line. The gene can be used to generate
 CC fusion proteins by linking to the gene to a human immunoglobulin FC
 CC portion (e.g. X79002) for increasing the stability of the fused protein
 CC as compared to the human protein only.
 CC The invention relates to 53 novel genes and their fragments (nucleic
 CC acid sequences: X79011-X79064; amino acid sequences Y14411-Y14464) which
 CC are useful for preventing, treating or ameliorating medical conditions
 CC e.g. by protein or gene therapy. Also, pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new polynucleotides.
 CC Specific uses are described for each of the 53 polynucleotides, based on
 CC which tissues they are most highly expressed in (see X79011 for described
 CC uses).

XX Sequence 42 AA;

Query Match 44.6%; Score 41; DB 20; Length 42;
 Best Local Similarity 60.0%; Pred. No. 10;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 CEWSAFSCFQ 12

Db 15 ckyshyscfq 24
 |::|::|::|

RESULT 9

W75033
 ID W75033 standard; Protein; 57 AA.

XX W75033;

XX 25-JAN-1999 (first entry)

DE Fragment of human secreted protein encoded by gene 64.

XX Human; secreted protein; testis; tumour; foetal brain tissue;
 KW fusion protein; cancer; central nervous system; seizure;
 KW diagnosis; neurodegenerative disease.

OS Homo sapiens.

XX WO9839448-A2.

XX 11-SEP-1998.

XX 06-MAR-1998; 98WO-US04493.

XX 02-OCT-1997; 97US-0061060.

XX 07-MAR-1997; 97US-0038621.

XX 07-MAR-1997; 97US-0040161.

XX 07-MAR-1997; 97US-0040162.

XX 07-MAR-1997; 97US-0040163.

XX 07-MAR-1997; 97US-0040333.

XX 07-MAR-1997; 97US-0040334.

XX 07-MAR-1997; 97US-0040336.

XX 07-MAR-1997; 97US-0040626.

XX 11-APR-1997; 97US-0043311.

XX 11-APR-1997; 97US-0043312.

```

PR 11-APR-1997; 97US-0043313.
PR 11-APR-1997; 97US-0043314.
PR 11-APR-1997; 97US-0043568.
PR 11-APR-1997; 97US-0043569.
PR 11-APR-1997; 97US-0043576.
PR 11-APR-1997; 97US-0043578.
PR 11-APR-1997; 97US-0043580.
PR 11-APR-1997; 97US-0043669.
PR 11-APR-1997; 97US-0043670.
PR 11-APR-1997; 97US-0043671.
PR 11-APR-1997; 97US-0043672.
PR 11-APR-1997; 97US-0043674.
PR 23-MAY-1997; 97US-0047492.
PR 23-MAY-1997; 97US-0047500.
PR 23-MAY-1997; 97US-0047501.
PR 23-MAY-1997; 97US-0047502.
PR 23-MAY-1997; 97US-0047503.
PR 23-MAY-1997; 97US-0047581.
PR 23-MAY-1997; 97US-0047582.
PR 23-MAY-1997; 97US-0047583.
PR 23-MAY-1997; 97US-0047584.
PR 23-MAY-1997; 97US-0047585.
PR 23-MAY-1997; 97US-0047586.
PR 23-MAY-1997; 97US-0047587.
PR 23-MAY-1997; 97US-0047588.
PR 23-MAY-1997; 97US-0047589.
PR 23-MAY-1997; 97US-0047590.
PR 23-MAY-1997; 97US-0047592.
PR 23-MAY-1997; 97US-0047593.
PR 23-MAY-1997; 97US-0047594.
PR 23-MAY-1997; 97US-0047595.
PR 23-MAY-1997; 97US-0047596.
PR 23-MAY-1997; 97US-0047597.
PR 23-MAY-1997; 97US-0047598.
PR 23-MAY-1997; 97US-0047599.
PR 23-MAY-1997; 97US-0047600.
PR 23-MAY-1997; 97US-0047601.
PR 23-MAY-1997; 97US-0047612.
PR 23-MAY-1997; 97US-0047613.
PR 23-MAY-1997; 97US-0047614.
PR 23-MAY-1997; 97US-0047615.
PR 23-MAY-1997; 97US-0047617.
PR 23-MAY-1997; 97US-0047618.
PR 23-MAY-1997; 97US-0047632.
PR 23-MAY-1997; 97US-0047633.
PR 06-JUN-1997; 97US-0048964.
PR 06-JUN-1997; 97US-0048974.
PR 13-JUN-1997; 97US-0049610.
PR 08-JUL-1997; 97US-0051926.
PR 16-JUL-1997; 97US-0052874.
PR 18-AUG-1997; 97US-0055724.
PR 22-AUG-1997; 97US-0056630.
PR 22-AUG-1997; 97US-0056631.
PR 22-AUG-1997; 97US-0056632.
PR 22-AUG-1997; 97US-0056636.
PR 22-AUG-1997; 97US-0056637.
PR 22-AUG-1997; 97US-0056662.
PR 22-AUG-1997; 97US-0056664.
PR 22-AUG-1997; 97US-0056845.
PR 22-AUG-1997; 97US-0056862.
PR 22-AUG-1997; 97US-0056864.
PR 22-AUG-1997; 97US-0056872.
PR 22-AUG-1997; 97US-0056874.
PR 22-AUG-1997; 97US-0056875.
PR 22-AUG-1997; 97US-0056876.
PR 22-AUG-1997; 97US-0056877.
PR 22-AUG-1997; 97US-0056878.
PR 22-AUG-1997; 97US-0056879.
PR 22-AUG-1997; 97US-0056880.
PR 22-AUG-1997; 97US-0056881.
PR 22-AUG-1997; 97US-0056882.
PR 22-AUG-1997; 97US-0056884.
PR 22-AUG-1997; 97US-0056886.

PR 22-AUG-1997; 97US-0056887.
PR 22-AUG-1997; 97US-0056888.
PR 22-AUG-1997; 97US-0056889.
PR 22-AUG-1997; 97US-0056892.
PR 22-AUG-1997; 97US-0056893.
PR 22-AUG-1997; 97US-0056894.
PR 22-AUG-1997; 97US-0056903.
PR 22-AUG-1997; 97US-0056908.
PR 22-AUG-1997; 97US-0056909.
PR 22-AUG-1997; 97US-0056910.
PR 05-SEP-1997; 97US-0056911.
PR 05-SEP-1997; 97US-0057650.
PR 05-SEP-1997; 97US-0057669.
PR 05-SEP-1997; 97US-0057761.
PR 12-SEP-1997; 97US-0058785.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PI Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
PI Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;
PI Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
XX WPI; 1998-506364/43.
XX
XX New isolated human genes and the secreted polypeptide(s) they encode
PT - useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
PS Disclosure; Page 43; 721pp; English.
XX
CC This sequence represents a fragment of a secreted human protein encoded
CC by the nucleic acid molecule designated Gene 64 (V59574).
CC The gene can be used to generate fusion proteins by linking to the gene
CC to a human immunoglobulin Fc portion (e.g. V59502) for increasing the
CC stability of the fused protein as compared to the human protein only.
CC The invention relates to 186 novel genes and their fragments (nucleic
CC acid sequences: V59511-V59812; amino acid sequences W74731-W75026) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 186 polynucleotides, based on
CC which tissues they are most highly expressed in (see V59511 for described
CC uses).
XX
SQ Sequence 57 AA;

Query Match 44.6%; Score 41; DB 19; Length 57;
Best Local Similarity 57.1%; Pred. No. 14;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 CEWSAFSCFQKAQL 16
  ||  || || || |
Db 3 ceicgftcrqkasl 16

RESULT 10
Y35840
ID Y35840 standard; Protein; 81 AA.
XX
AC Y35840;
XX
DT 13-SEP-1999 (first entry)
XX
DE Chlamydia pneumoniae lipoprotein sequence.
XX
KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope.
XX
OS Chlamydia pneumoniae.

```


XX	11-SEP-1998.	98WO-US04493.
PD		
XX		
XX		
PF	06-MAR-1998;	
XX		
XX	02-OCT-1997;	97US-0061060.
PR	07-MAR-1997;	97US-0038621.
PR	07-MAR-1997;	97US-0040161.
PR	07-MAR-1997;	97US-0040162.
PR	07-MAR-1997;	97US-0040163.
PR	07-MAR-1997;	97US-0040333.
PR	07-MAR-1997;	97US-0040334.
PR	07-MAR-1997;	97US-0040336.
PR	07-MAR-1997;	97US-0040626.
PR	11-APR-1997;	97US-0043311.
PR	11-APR-1997;	97US-0043312.
PR	11-APR-1997;	97US-0043313.
PR	11-APR-1997;	97US-0043311.
PR	11-APR-1997;	97US-0043566.
PR	11-APR-1997;	97US-0043569.
PR	11-APR-1997;	97US-0043576.
PR	11-APR-1997;	97US-0043578.
PR	11-APR-1997;	97US-0043580.
PR	11-APR-1997;	97US-0043669.
PR	11-APR-1997;	97US-0043670.
PR	11-APR-1997;	97US-0043671.
PR	11-APR-1997;	97US-0043672.
PR	11-APR-1997;	97US-0043674.
PR	23-MAY-1997;	97US-0047492.
PR	23-MAY-1997;	97US-0047500.
PR	23-MAY-1997;	97US-0047501.
PR	23-MAY-1997;	97US-0047502.
PR	23-MAY-1997;	97US-0047503.
PR	23-MAY-1997;	97US-0047581.
PR	23-MAY-1997;	97US-0047582.
PR	23-MAY-1997;	97US-0047583.
PR	23-MAY-1997;	97US-0047584.
PR	23-MAY-1997;	97US-0047585.
PR	23-MAY-1997;	97US-0047586.
PR	23-MAY-1997;	97US-0047587.
PR	23-MAY-1997;	97US-0047588.
PR	23-MAY-1997;	97US-0047589.
PR	23-MAY-1997;	97US-0047590.
PR	23-MAY-1997;	97US-0047592.
PR	23-MAY-1997;	97US-0047593.
PR	23-MAY-1997;	97US-0047594.
PR	23-MAY-1997;	97US-0047595.
PR	23-MAY-1997;	97US-0047596.
PR	23-MAY-1997;	97US-0047597.
PR	23-MAY-1997;	97US-0047598.
PR	23-MAY-1997;	97US-0047599.
PR	23-MAY-1997;	97US-0047600.
PR	23-MAY-1997;	97US-0047601.
PR	23-MAY-1997;	97US-0047612.
PR	23-MAY-1997;	97US-0047613.
PR	23-MAY-1997;	97US-0047614.
PR	23-MAY-1997;	97US-0047615.
PR	23-MAY-1997;	97US-0047617.
PR	23-MAY-1997;	97US-0047618.
PR	23-MAY-1997;	97US-0047632.
PR	23-MAY-1997;	97US-0047633.
PR	06-JUN-1997;	97US-0048964.
PR	06-JUN-1997;	97US-0048974.
PR	13-JUN-1997;	97US-0049610.
PR	16-JUL-1997;	97US-0051926.
PR	18-AUG-1997;	97US-0052874.
PR	18-AUG-1997;	97US-0055724.
PR	22-AUG-1997;	97US-0056630.
PR	22-AUG-1997;	97US-0056631.
PR	22-AUG-1997;	97US-0056632.
PR	22-AUG-1997;	97US-0056636.
PR	22-AUG-1997;	97US-0056637.
PR	22-AUG-1997;	97US-0056662.

PR	22-AUG-1997;	97US-0056664.	RESULT 12
PR	22-AUG-1997;	97US-0056845.	Y73348
PR	22-AUG-1997;	97US-0056862.	ID Y73348 standard; Protein; 241 AA.
PR	22-AUG-1997;	97US-0056884.	XX
PR	22-AUG-1997;	97US-0056872.	XX Y73348;
PR	22-AUG-1997;	97US-0056874.	XX
PR	22-AUG-1997;	97US-0056875.	DT 24-FEB-2000 (first entry)
PR	22-AUG-1997;	97US-0056876.	XX
PR	22-AUG-1997;	97US-0056877.	XX
PR	22-AUG-1997;	97US-0056878.	DE HTRM clone 839651 protein sequence.
PR	22-AUG-1997;	97US-0056879.	XX
PR	22-AUG-1997;	97US-0056880.	KW HTRM: human transcriptional regulatory molecule; arteriosclerosis; AIDS;
PR	22-AUG-1997;	97US-0056881.	KW arteriosclerosis; cirrhosis; cancer; leukaemia; diabetes mellitus;
PR	22-AUG-1997;	97US-0056882.	KW Addison's disease; multiple sclerosis; rheumatoid arthritis; infection;
PR	22-AUG-1997;	97US-0056884.	KW trauma; myasthenia gravis; adenocarcinoma; immune disorder; treatment.
PR	22-AUG-1997;	97US-0056886.	XX
PR	22-AUG-1997;	97US-0056887.	OS Homo sapiens.
PR	22-AUG-1997;	97US-0056888.	XX
PR	22-AUG-1997;	97US-0056889.	PN WO9957144-A2.
PR	22-AUG-1997;	97US-0056892.	XX
PR	22-AUG-1997;	97US-0056893.	PD 11-NOV-1999.
PR	22-AUG-1997;	97US-0056894.	XX
PR	22-AUG-1997;	97US-0056903.	PF 04-MAY-1999; 99WO-US09935.
PR	22-AUG-1997;	97US-0056908.	XX
PR	22-AUG-1997;	97US-0056909.	PR 05-MAY-1998; 98US-0084254.
PR	22-AUG-1997;	97US-0056910.	PR 07-AUG-1998; 98US-0095827.
PR	22-AUG-1997;	97US-0056911.	PR 02-OCT-1998; 98US-0102745.
PR	05-SEP-1997;	97US-0057650.	XX (INCY-) INCYTE PHARM INC.
PR	05-SEP-1997;	97US-0057669.	XX
PR	05-SEP-1997;	97US-0057761.	PI Hillman JL, Bandman O, Lal P, Yue H, Reddy R, Tang YT;
PR	12-SEP-1997;	97US-0058785.	PI Gerstin EH, Patterson C, Baughn MR, Azimzai Y, Lu DAM;
XX			DR WPI; 2000-052941/04.
XX	(HUMA-) HUMAN GENOME SCI INC.		DR N-PSDB; Z52433.
XX			XX
PI	Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;		XX
PI	Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;		XX
PI	Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;		PT New peptides useful for diagnosis, prevention and treatment of cancer
PI	Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;		PT and immune disorders
XX			XX
DR	WPI; 1998-506364/43.		PS Claim 1; Page 109; 193pp; English.
XX			XX
PT	New isolated human genes and the secreted polypeptide(s) they encode		CC Y73325-Y73389 are human transcriptional regulator molecule (HTRM) protein
PT	- useful for diagnosis and treatment of e.g. cancers, neurological		CC sequences. The HTRM protein and nucleotide sequences are useful for
PT	disorders, immune diseases, inflammation or blood disorders		CC preventing or treating disorders associated with decreased expression or
XX			CC activity of HTRM which include cell proliferative disorders such as
XX	Disclosure; Page 43; 721pp; English.		CC arteriosclerosis and cirrhosis; cancers including adenocarcinoma and
XX			CC leukaemia; immune disorders such as AIDS, Addison's disease, diabetes
XX			CC mellitus, rheumatoid arthritis, multiple sclerosis, systemic lupus
CC	This sequence represents a fragment of a secreted human protein encoded		CC erythematous, and myasthenia gravis; infections and trauma. Antagonists
CC	by the nucleic acid molecule designated Gene 64 (V59574).		CC of the HTRM polypeptides are useful for treating or preventing disorders
CC	The gene can be used to generate fusion proteins by linking to the gene		CC associated with increased expression or activity of HTRMs. HTRM
CC	to a human immunoglobulin Fc portion (e.g. V59502) for increasing the		CC polypeptides, their immunogenic fragments or oligopeptides are useful for
CC	stability of the fused protein as compared to the human protein only.		CC screening libraries of compounds in drug screening techniques.
CC	The invention relates to 186 novel genes and their fragments (nucleic		CC Polynucleotides encoding HTRM are useful for blocking the transcription
CC	acid sequences: V59511-V59812; amino acid sequences W74731-W75026) which		CC of mRNA and regulating gene function by modulating the activity of HTRM.
CC	are useful for preventing, treating or ameliorating medical conditions		CC Vectors expressing HTRM or agonists can also be used to prevent or treat
CC	e.g. by protein or gene therapy. Also, pathological conditions can be		CC disorder associated with decreased HTRM expression. Antibodies which
CC	diagnosed by determining the amount of the new polypeptides in a sample		CC specifically bind HTRM and polynucleotides encoding HTRM are useful for
CC	or by determining the presence of mutations in the new polynucleotides.		CC diagnosing disorders associated with the expression of HTRM, particularly
CC	Specific uses are described for each of the 186 polynucleotides, based on		CC in assays that detect the expression of HTRM. Nucleotide sequences
CC	which tissues they are most highly expressed in (see V59511 for described		CC encoding HTRM may be useful to generate hybridization probes useful in
XX	uses).		CC mapping the naturally occurring genomic sequence and to detect
XX			CC differences in gene sequences among normal, carrier and affected
XX			CC individuals. Using diagnostic assays, cancer can be detected prior to the
XX			CC appearance of clinical symptoms and thereby progression of cancer can be
XX			CC prevented by aggressive treatment or preventive measures.
SQ	Sequence 183 AA;		SQ Sequence 241 AA;
Query Match	44.6%; Score 41; DB 19; Length 183;		Query Match 44.6%; Score 41; DB 21; Length 241;
Best Local Similarity	57.1%; Pred. No. 44;		Best Local Similarity 57.1%; Pred. No. 58;
Matches	8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;		Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Oy	3 CEWSAFSCFOKAOL 16		
Db	3 ceigqfctcrqasl 16		

QY 3 CEWSAFSCFOKAOL 16
 || | | | |
 Db 141 ceicgftcrqkasi 154

RESULT 13
 W36002
 ID W36002 standard; Protein; 570 AA.

XX AC W36002;
 XX 03-MAR-1998 (first entry)
 DT Human Fchd531 gene product.
 DE Fchd531 gene; differential expression; endothelial cell; human;
 KW shear stress; cardiovascular disease; atherosclerosis; ischaemia;
 KW reperfusion; hypertension; restenosis; arterial inflammation;
 KW therapy; diagnosis; drug screening; marker.

XX OS Homo sapiens.
 XX PN WO9730065-A1.
 XX PD 21-AUG-1997.

XX PF 14-FEB-1997; 97WO-US02291.
 XX PR 13-FEB-1997; 97US-0799910.
 XX PR 16-FEB-1996; 96US-0011787.

XX PA (MILL-) MILLENNIUM PHARM INC.
 XX PI Falb DA;
 XX DR WPI; 1997-424966/39.
 XX N-PSDB; T94467.

XX New genes differentially expressed in cardiovascular disease - used
 PT for diagnosis, drug screening and treatment of cardiovascular
 PT disease, e.g. atherosclerosis, restenosis, hypertension, etc
 PS Example 7; Fig 1A-1B; 163pp; English.

XX This protein is encoded by the novel human fchd531 gene (see
 CC T94467) that is down-regulated in endothelial cells subjected to
 CC turbulent and laminar shear stress. Shear stress is thought to be
 CC responsible for the prevalence of atherosclerotic lesions in areas
 CC of unusual circulatory flow. The fchd531 gene product has 94%
 CC similarity to the mouse penta zinc finger gene (Pzf). Gene
 CC products were also identified for novel genes fchd540 (see
 CC W36003) and fchd545 (see W36004), which are respectively up- and
 CC down-regulated in endothelial cells subjected to shear stress.
 CC Novel fchd531, fchd540, fchd545, fchd602 and fchd605 genes (see
 CC T94467-71) provide a fingerprint for the study of cardiovascular
 CC diseases, including atherosclerosis, ischaemia/reperfusion,
 CC hypertension, restenosis and arterial inflammation. Methods are
 CC provided for the diagnosis, monitoring in clinical trials,
 CC screening for therapeutically effective compounds, and treatment of
 CC cardiovascular diseases based on discoveries regarding the
 CC expression patterns of these novel genes.

XX Sequence 570 AA;

Query Match 44.6%; Score 41; DB 18; Length 570;
 Best Local Similarity 57.1%; Pred. No. 1.4e+02;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 CEWSAFSCFOKAOL 16
 || | | | |
 Db 402 ceicgftcrqkasi 415

RESULT 14
 Y45013
 ID Y45013 standard; Protein; 570 AA.

XX AC Y45013;
 XX 31-MAY-2000 (first entry)
 DT Protein encoded by fchd531 gene.

XX KW fchd531 gene; human; cardiovascular disease; oncogenic disorder;
 KW diabetic retinopathy; fibroproliferative disorder; atherosclerosis;
 KW TGF-beta signalling pathway; TGF; Transforming growth factor;
 KW pancreatic cancer; angiogenesis; inflammation; fibrosis; tumour growth;
 KW vascularisation; cytostatic; antidiabetic; ophthalmological.

XX OS Homo sapiens.
 XX PN WO200006206-A1.
 XX PD 10-FEB-2000.

XX PF 30-JUL-1999; 99WO-US17394.
 XX PR 30-JUL-1998; 98US-0126640.

XX PA (MILL-) MILLENNIUM PHARM INC.
 XX PI Falb DA;
 XX DR WPI; 2000-205414/18.
 XX N-PSDB; 250707.

XX Identifying substances for ameliorating symptoms of fibroproliferative
 XX diseases or oncogenic related disorders -

XX Example; Fig 1; 214pp; English.

XX The patent discloses methods for the treatment and diagnosis of
 CC cardiovascular diseases by novel human genes which are differentially
 CC expressed in different cardiovascular disease states. Compositions which
 CC can modify TGF-beta signalling pathway are identified by screening.
 CC These are used therapeutically to treat fibroproliferative and oncogenic
 CC disorders, especially TGF (Transforming growth factor)-beta related
 CC disorders, including diabetic retinopathy, atherosclerosis, pancreatic
 CC cancer, angiogenesis, inflammation, fibrosis, tumour growth and
 CC vascularisation. The present sequence is the protein product of fchd531
 CC gene which is down-regulated in endothelial cells subjected to shear
 CC stress can be used to design cardiovascular disease treatment
 CC strategies. Depending on whether the down-regulation has a pathogenic or
 CC protective effect treatment methods can be designed to increase or
 CC decrease the activity of the protein product of the gene.

XX Sequence 570 AA;

Query Match 44.6%; Score 41; DB 21; Length 570;
 Best Local Similarity 57.1%; Pred. No. 1.4e+02;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 CEWSAFSCFOKAOL 16
 || | | | |
 Db 402 ceicgftcrqkasi 415

RESULT 15
 R91428
 ID R91428 standard; Protein; 1111 AA.

XX AC R91428;
 XX

DT 13-NOV-1996 (first entry)
 XX
 DE Kalinin/laminin 5 gamma-2 chain (alternative form).
 XX
 KW kalinin; laminin; epidermolysis bullosa; junctional; probe;
 XX detection; inhibit; monitor; malignancy.
 KW
 OS Homo sapiens.
 XX
 PN WO9610646-A1.
 XX
 XX 11-APR-1996.
 XX
 PF 04-OCT-1995; 95WO-EP03918.
 XX
 XX 04-OCT-1994; 94US-0317450.
 XX
 XX (TRYG/) TRYGGVASON K.
 XX
 PI Kallunki P, Pyke C, Tryggvason K;
 XX
 XX WPI; 1996-209366/21.
 DR N-PSDB; T13323.
 DR
 XX
 PT Detection of kalinin or laminin 5 expression in cells - useful to
 PT detect, monitor and inhibit the invasive growth of cell in tissue,
 PT partic. malignant tissue
 XX
 PS Disclosure; Fig 4B; 37pp; English.
 XX
 CC The present sequence is an alternative form of kalinin/laminin 5 gamma-2
 CC chain (see R91427). The gamma-2 chain is of importance to patients
 CC suffering from epidermolysis bullosa, esp. the junctional form (JEB).
 CC Probes and antisense gamma-2 sequences derived from this sequence can be
 CC used to detect, monitor and inhibit the invasive growth of cells in
 CC tissue, partic. malignant tissue.
 CC
 XX
 SQ Sequence 1111 AA;

Query Match 44.6%; Score 41; DB 17; Length 1111;
 Best Local Similarity 77.8%; Pred. No. 2.6e+02;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 NCEWSAFSC 10
 ||| ||||
 Db 601 ncehgafsc 609

Search completed: May 23, 2001, 11:11:37
 Job time: 180 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:09:43 ; Search time 58.85 Seconds
(without alignments)
5.223 Million cell updates/sec

Title: US-09-522-217-2_COPY_69_84

Perfect score: 92

Sequence: 1 TNCWSAFSCFQKAQL 16

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2.6/plodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2.6/plodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2.6/plodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2.6/plodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2.6/plodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2.6/plodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	44.6	570	3	US-08-826-246-2
2	41	44.6	570	3	US-08-944-495-2
3	41	44.6	570	3	US-09-126-640-7
4	41	44.6	1111	1	US-08-317-450B-15
5	41	44.6	1111	4	US-08-800-593-15
6	41	44.6	1193	4	US-08-317-450B-13
7	41	44.6	1193	4	US-08-800-593-13
8	39	42.4	124	1	US-08-170-360-1
9	39	42.4	126	1	US-08-170-360-3
10	39	42.4	126	2	US-08-888-497-38
11	39	42.4	126	5	PCT-US94-07926-38
12	39	42.4	148	2	US-08-888-497-36
13	39	42.4	148	5	PCT-US94-07926-36
14	39	42.4	536	1	US-08-164-614A-12
15	39	42.4	536	2	US-08-456-489B-12
16	39	42.4	1266	1	US-08-468-557-4
17	38	41.3	214	1	US-08-766-605-1
18	38	41.3	214	2	US-09-094-212-1
19	38	41.3	263	2	US-08-634-924B-2
20	38	41.3	522	1	US-08-639-237-2
21	38	41.3	522	1	US-08-975-405-2
22	37	40.2	418	2	US-08-290-731C-13
23	36	39.1	72	1	US-08-379-538-1
24	36	39.1	72	1	US-08-379-538-6
25	36	39.1	269	2	US-07-857-224B-48
26	36	39.1	271	2	US-07-857-224B-47
27	36	39.1	307	1	US-08-164-614A-11

28	36	39.1	307	2	US-08-456-489B-11	Sequence 11, Appl
29	36	39.1	522	1	US-08-164-614A-10	Sequence 10, Appl
30	36	39.1	522	1	US-08-456-489B-10	Sequence 10, Appl
31	36	39.1	717	4	US-08-872-855-9	Sequence 9, Appl
32	36	39.1	721	4	US-08-872-855-7	Sequence 7, Appl
33	36	39.1	729	4	US-08-872-855-8	Sequence 8, Appl
34	36	39.1	933	2	US-08-313-200-1	Sequence 1, Appl
35	36	39.1	933	5	PCT-US93-03837-1	Sequence 1, Appl
36	36	39.1	1048	4	US-09-356-952-5	Sequence 5, Appl
37	36	39.1	1139	1	US-08-537-210A-4	Sequence 4, Appl
38	36	39.1	1139	4	US-09-113-825-4	Sequence 4, Appl
39	36	39.1	2703	1	US-08-185-432-19	Sequence 19, Appl
40	35	38.0	118	3	US-09-090-602-4	Sequence 4, Appl
41	35	38.0	559	2	US-08-884-072-6	Sequence 6, Appl
42	35	38.0	788	2	US-08-918-914-4	Sequence 4, Appl
43	35	38.0	908	2	US-08-588-526-3	Sequence 3, Appl
44	35	38.0	981	2	US-08-649-046-2	Sequence 2, Appl
45	35	38.0	1014	4	US-09-078-347A-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-826-246-2
; Sequence 2, Application US/08826246
; Patent No. 6048709
; GENERAL INFORMATION:
; APPLICANT: Falb, Dean
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE TREATMENT AND DIAGNOSIS OF
; NUMBER OF SEQUENCE: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,246
; FILING DATE: 28-MAR-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/799,910
; FILING DATE: 13-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/011,787
; FILING DATE: 16-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-078-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)7909090
; TELEFAX: (212)8699741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 570 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-826-246-2

Query Match 44.6%; Score 41; DB 3; Length 570;
Best Local Similarity 57.1%; Pred. No. 40;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

RESULT 2

US-08-944-495-2
; Sequence 2, Application US/08944495
; Patent No. 6087477
; GENERAL INFORMATION:
; APPLICANT: Falb, Dean
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE TREATMENT AND DIAGNOSIS OF
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; City: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; APPLICATION NUMBER: US/08/944.495
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/799,910
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-067-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)7909090
; TELEFAX: (212)8699741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 570 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal

US-08-944-495-2
Query Match 44.6%; Score 41; DB 3; Length 570;
Best Local Similarity 57.1%; Pred. No. 40;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

RESULT 3

US-09-126-640-7
; Sequence 7, Application US/09126640A
; Patent No. 6099823
; GENERAL INFORMATION:
; APPLICANT: Falb, Dean A.

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; FILE REFERENCE: 7853-126
; CURRENT APPLICATION NUMBER: US/09/126,640A
; CURRENT FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 08/870,434
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 08/799,910
; EARLIER FILING DATE: 1997-02-13
; EARLIER APPLICATION NUMBER: 60/011,787
; EARLIER FILING DATE: 1996-02-16
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 570
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-126-640-7

Query Match 44.6%; Score 41; DB 3; Length 570;
Best Local Similarity 57.1%; Pred. No. 40;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Qy 3 CEWSAFSCFQKAO 16
|| | | | | | |
Db 402 CEICGFTCQKASL 415

Best Local Similarity 77.8%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 NCEWSAFSC 10
Db 601 NCEHGAFC 609

RESULT 5
US-08-800-593-15
; Sequence 15, Application US/08800593
; Patent No. 6143505
; GENERAL INFORMATION:
; APPLICANT: Tryggvason, Karl
; APPLICANT: Kallunki, Pekka
; APPLICANT: Pyke, Charles
; TITLE OF INVENTION: Laminin Chains: Diagnostic and
; TITLE OF INVENTION: Therapeutic Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 18-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 08/317,450
; APPLICATION NUMBER: US 08/317,450
; FILING DATE: 04-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chao, Mark
; REGISTRATION NUMBER: 37,293
; REFERENCE/DOCKET NUMBER: 94,778-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1111 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-800-593-15

Query Match 44.6%; Score 41; DB 4; Length 1111;
Best Local Similarity 77.8%; Pred. No. 79;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 NCEWSAFSC 10
Db 601 NCEHGAFC 609

RESULT 6
US-08-317-450B-13
; Sequence 13, Application US/08317450B
; Patent No. 5660982
; GENERAL INFORMATION:
; APPLICANT: Tryggvason, Karl
; APPLICANT: Kallunki, Pekka
; APPLICANT: Pyke, Charles
; TITLE OF INVENTION: Laminin Chains: Diagnostic and

; TITLE OF INVENTION: Therapeutic Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER & ALLEGRETTI, LTD.
; STREET: Ten South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 04-OCT-1994
; APPLICATION NUMBER: US/08/317,450B
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chao, Mark
; REGISTRATION NUMBER: 37,293
; REFERENCE/DOCKET NUMBER: 94,778
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1193 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-317-450B-13

Query Match 44.6%; Score 41; DB 1; Length 1193;
Best Local Similarity 77.8%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 NCEWSAFSC 10
Db 601 NCEHGAFC 609

RESULT 7
US-08-800-593-13
; Sequence 13, Application US/08800593
; Patent No. 6143505
; GENERAL INFORMATION:
; APPLICANT: Tryggvason, Karl
; APPLICANT: Kallunki, Pekka
; APPLICANT: Pyke, Charles
; TITLE OF INVENTION: Laminin Chains: Diagnostic and
; TITLE OF INVENTION: Therapeutic Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,593
; FILING DATE: 18-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/317,450
; FILING DATE: 04-OCT-1994

```
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chao, Mark
; REGISTRATION NUMBER: 37,293
; REFERENCE/DOCKET NUMBER: 94,778-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1193 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-800-593-13

Query Match 44.6%; Score 41; DB 4; Length 1193;
Best Local Similarity 77.8%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 NCWSAFSC 10
   III IIII
DB 601 NCEHGAFC 609

RESULT 8
US-08-170-360-1
; Sequence 1, Application US/08170360
; Patent No. 5656602
; GENERAL INFORMATION:
; APPLICANT: Tseng, Albert P. S.
; APPLICANT: Ingilis, Adam
; TITLE OF INVENTION: PLA2 INHIBITORY COMPOUNDS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg Ernst & Kurz
; STREET: Suite 701-E, 555 Thirteenth St., N.W
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/170,360
; FILING DATE: 03-MAR-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PCT/AU92/00333
; FILING DATE: 06-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PK 7058
; FILING DATE: 04-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Ernst, Barbara G.
; REGISTRATION NUMBER: 30,377
; REFERENCE/DOCKET NUMBER: 1871-104A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)783-6040
; TELEFAX: (202)783-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-170-360-3

Query Match 42.4%; Score 39; DB 1; Length 124;
Best Local Similarity 53.8%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 NCEWSAFSCFOKA 14
   II: :I II II
DB 97 NCDRNAICFSKA 109

US-08-170-360-3
; Sequence 3, Application US/08170360
; Patent No. 5656602
; GENERAL INFORMATION:
; APPLICANT: Tseng, Albert P. S.
; APPLICANT: Ingilis, Adam
; TITLE OF INVENTION: PLA2 INHIBITORY COMPOUNDS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg Ernst & Kurz
; STREET: Suite 701-E, 555 Thirteenth St., N.W
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/170,360
; FILING DATE: 03-MAR-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PCT/AU92/00333
; FILING DATE: 06-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PK 7058
; FILING DATE: 04-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Ernst, Barbara G.
; REGISTRATION NUMBER: 30,377
; REFERENCE/DOCKET NUMBER: 1871-104A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)783-6040
; TELEFAX: (202)783-6031
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-170-360-3

Query Match 42.4%; Score 39; DB 1; Length 126;
Best Local Similarity 53.8%; Pred. No. 18;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 NCEWSAFSCFOKA 14
   II: :I II II
DB 97 NCDRNAICFSKA 109
```



```

RESULT 10
US-08-888-497-38
; Sequence 38, Application US/08888497
; Patent No. 5972677
; GENERAL INFORMATION:
; APPLICANT: Tischfield, Jay A.
; TITLE OF INVENTION: Mammalian Phospholipase A2 Nucleotide
; TITLE OF INVENTION: Sequences and Low Molecular Weight Amino Acid Sequences
; TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide
; TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &
; ADDRESSEE: Russell PA
; STREET: 200 East Broward Boulevard
; CITY: Fort Lauderdale
; STATE: FL
; COUNTRY: USA
; ZIP: 33301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/888,497
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/651,405
; FILING DATE:
; APPLICATION NUMBER: US 08/097,354
; FILING DATE: 26-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Manso, Peter J.
; REGISTRATION NUMBER: 32,264
; REFERENCE/DOCKET NUMBER: IN21044-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 305-527-2498
; TELEFAX: 305-764-4996
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-07926-38
; Query Match 42.4%; Score 39; DB 2; Length 126;
; Best Local Similarity 53.8%; Pred. No. 18;
; Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 NCSEAFSCFQKA 14
Db 97 NCDRNAICFCSKA 109
||:| ||||

RESULT 11
US-08-888-497-38
; Sequence 38, Application PC/TUS9407926
; Patent No. 5972677
; GENERAL INFORMATION:
; APPLICANT: Tischfield, Jay A.
; TITLE OF INVENTION: Mammalian Phospholipase A2 Nucleotide
; TITLE OF INVENTION: Sequences and Low Molecular Weight Amino Acid Sequences
; TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide
; TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &
; ADDRESSEE: Russell PA
; STREET: 200 East Broward Boulevard
; CITY: Fort Lauderdale
; STATE: FL
; COUNTRY: USA
; ZIP: 33301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/888,497
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/651,405
; FILING DATE:
; APPLICATION NUMBER: US 08/097,354
; FILING DATE: 26-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Manso, Peter J.
; REGISTRATION NUMBER: 32,264
; REFERENCE/DOCKET NUMBER: IN21044-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 305-527-2498
; TELEFAX: 305-764-4996
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-07926-38
; Query Match 42.4%; Score 39; DB 2; Length 126;
; Best Local Similarity 53.8%; Pred. No. 18;
; Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 NCSEAFSCFQKA 14
Db 97 NCDRNAICFCSKA 109
||:| ||||

RESULT 12
US-08-888-497-36
; Sequence 36, Application US/08888497
; Patent No. 5972677
; GENERAL INFORMATION:
; APPLICANT: Tischfield, Jay A.
; TITLE OF INVENTION: Mammalian Phospholipase A2 Nucleotide
; TITLE OF INVENTION: Sequences and Low Molecular Weight Amino Acid Sequences
; TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide
; TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &
; ADDRESSEE: Russell PA
; STREET: 200 East Broward Boulevard
; CITY: Fort Lauderdale
; STATE: FL
; COUNTRY: USA
; ZIP: 33301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/888,497
; FILING DATE:
; CLASSIFICATION:

```

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/651,405
;; FILING DATE: 26-JUL-1993
;; APPLICATION NUMBER: US 08/097,354
;; FILING DATE: 26-JUL-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Manso, Peter J.
;; REGISTRATION NUMBER: 32,264
;; REFERENCE/DOCKET NUMBER: IN21044-5
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 305-527-2498
;; TELEFAX: 305-764-4996
;; INFORMATION FOR SEQ ID NO: 36:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 148 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-888-497-36

Query Match 42.4%; Score 39; DB 2; Length 148;
Best Local Similarity 53.8%; Pred. No. 21;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy 2 NCEWSAFSCFOKA 14
Db 119 NCDRNAAICFSKA 131

RESULT 13
PCT-US94-07926-36
;; Sequence 36, Application PC/TUS9407926
;; GENERAL INFORMATION:
;; APPLICANT: Tischfield, Jay A.
;; APPLICANT: Seilhamer, Jeffrey J.
;; TITLE OF INVENTION: Mammalian Phospholipase A2 Nucleotide
;; TITLE OF INVENTION: Sequences and Low Molecular Weight Amino Acid Sequences
;; TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide
;; TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites
;; NUMBER OF SEQUENCES: 44
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &
;; ADDRESSEE: Russell PA
;; STREET: 200 East Broward Boulevard
;; CITY: Fort Lauderdale
;; STATE: FL
;; COUNTRY: USA
;; ZIP: 33301
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US94/07926
;; FILING DATE: 15-JUL-1994
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/097,354
;; FILING DATE: 26-JUL-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Manso, Peter J.
;; REGISTRATION NUMBER: 32,264
;; REFERENCE/DOCKET NUMBER: IN21044-5
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 305-527-2498
;; TELEFAX: 305-764-4996
;; INFORMATION FOR SEQ ID NO: 36:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 148 amino acids
;; TYPE: amino acid

;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
PCT-US94-07926-36

Query Match 42.4%; Score 39; DB 5; Length 148;
Best Local Similarity 53.8%; Pred. No. 21;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy 2 NCEWSAFSCFOKA 14
Db 119 NCDRNAAICFSKA 131

RESULT 14
US-08-164-614A-12
;; Sequence 12, Application US/08164614A
;; Patent No. 5789237
;; GENERAL INFORMATION:
;; APPLICANT: Renauld, Jean-Christophe
;; APPLICANT: Druez, Catherine
;; APPLICANT: Van Snick, Jacques
;; TITLE OF INVENTION: Nucleic Acid Sequences Coding For
;; TITLE OF INVENTION: Or Complementary To Nucleic Acid Sequences Coding For
;; TITLE OF INVENTION: Interleukin 9 Receptor
;; NUMBER OF SEQUENCES: 12
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Felfe & Lynch
;; STREET: 805 Third Avenue
;; CITY: New York City
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10022
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
;; COMPUTER: IBM PS/2
;; OPERATING SYSTEM: PC-DOS
;; SOFTWARE: Wordperfect
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/164,614A
;; FILING DATE: 8-DECEMBER-1993
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/847,347
;; FILING DATE: 09-MARCH-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hanson, No. 5789237man D.
;; REGISTRATION NUMBER: 30,946
;; REFERENCE/DOCKET NUMBER: LUD 5264.1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 688-9200
;; TELEFAX: (212) 838-3884
;; INFORMATION FOR SEQ ID NO: 12:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 536 amino acid residues
;; TYPE: amino acid
;; TOPOLOGY: linear
US-08-164-614A-12

Query Match 42.4%; Score 39; DB 1; Length 536;
Best Local Similarity 58.3%; Pred. No. 78;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 4 EWSAFSCFQRAQ 15
Db 261 EWSQPVCFORPO 272

RESULT 15
US-08-456-489B-12
;; Sequence 12, Application US/08456489B

Patent No. 5962269
GENERAL INFORMATION:
APPLICANT: Renaud, Jean-Christophe; Druetz, Catherine; Van Snick,
APPLICANT: Jacques
TITLE OF INVENTION: Nucleic Acid Sequences Coding For Or
TITLE OF INVENTION: Complementary To Nucleic Acid Sequences Coding For Interleukin
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski L.L.P.
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,489B
FILING DATE: 1-JUNE-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/164,614
FILING DATE: 8-DECEMBER-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847,347
FILING DATE: 09-MARCH-1992
ATTORNEY/AGENT INFORMATION:
NAME: Schofield, Mary Anne
REGISTRATION NUMBER: 36,669
REFERENCE/DOCKET NUMBER: LUD 5264.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3000
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 536 amino acid residues
TYPE: amino acid
TOPOLOGY: linear
US-08-456-489B-12

Query Match 42.4%; Score 39; DB 2; Length 536;
Best Local Similarity 58.3%; Pred. No. 78;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 EWSAFSCFQKAO 15
Db 261 EWSQPVCFORPQ 272

Search completed: May 23, 2001, 11:09:43
Job time: 66 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:12:52 ; Search time 70.54 Seconds
(without alignments)
15.588 Million cell updates/sec

Title: US-09-522-217-2_COPY_69_84
Perfect score: 92
Sequence: 1 TNCWAFSCFQAQL 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_67.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	50.0	1173	2 T30608	proteophosphoglycan
2	45	48.9	1403	2 S64142	hypothetical prote
3	42	45.7	146	1 PSDG	phospholipase A2 (
4	42	45.7	441	1 TVCHMC	transforming prote
5	42	45.7	500	2 T51172	betaine-aldehyde d
6	41.5	45.1	368	2 S57273	lignin peroxidase
7	41	44.6	315	2 T42682	hypothetical prote
8	41	44.6	354	2 I48722	zinc finger protei
9	41	44.6	455	2 I48724	zinc finger protei
10	41	44.6	1193	2 A44018	laminin B2t chain
11	41	44.6	3344	2 J01899	genome polyprotein
12	41	44.6	4735	2 T17463	rifamycin polyketi
13	40	43.5	142	2 T48816	hypothetical prote
14	40	43.5	219	2 T37794	hypothetical prote
15	40	43.5	505	2 T03394	probable betaine-a
16	40	43.5	533	2 T33912	hypothetical prote
17	39	42.4	124	1 P5PGA2	phospholipase A2 (
18	39	42.4	132	1 P5HOA	phospholipase A2 (
19	39	42.4	146	1 P5PGA	phospholipase A2 (
20	39	42.4	146	1 S34049	phospholipase A2 (
21	39	42.4	148	1 PSHU	phospholipase A2 (
22	39	42.4	223	1 VCBVCA	coat protein - tob
23	39	42.4	806	2 T00470	XE169 protein homo
24	39	42.4	1146	2 T02766	adhesin WI-1 - Aje
25	39	42.4	1257	2 S44754	Cl489.8 protein -
26	39	42.4	1266	2 I59314	isoleucine--tRNA l
27	39	42.4	2395	1 S50820	surface protein ty
28	38.5	41.8	230	2 S50348	probable membrane
29	38	41.3	209	2 E59641	phosphoribosyl-AMP

Query Match 50.0%; Score 46; DB 2; Length 1173;
Best Local Similarity 60.0%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Query 3 CEWSAFSCFQ 12
Db 803 CEWATFLCFR 812
|||:| | |
| | | : | | |

RESULT 2
S64142
hypothetical protein YGL131c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein G2850
C;Species: Saccharomyces cerevisiae
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 21-Nov-1997
C;Accession: S64142; S64144
R;Cerdan, E.; Rodriguez-Torres, A.M.; Rodriguez-Belmonte, E.; Tizon, B.; Cadahia, J.L
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64134
A;Accession: S64142
A;Molecule type: DNA
A;Residues: 1-1398 <CER>
A;Cross-references: EMBL:272654; MIPS:YGL131c
A;Experimental source: strain S288C
R;Escribano, V.; Eraso, P.; Portillo, F.; Mazon, M.J.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64144
A;Accession: S64144
A;Molecule type: DNA
A;Residues: 1232-1403 <ESC>
A;Cross-references: EMBL:272654; MIPS:YGL131c
A;Experimental source: strain S288C

RESULT 1
T30608
proteophosphoglycan homolog 6L - Molluscum contagiosum virus 1
N;Alternate names: MC006L
C;Species: Molluscum contagiosum virus 1
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 11-May-2000
C;Accession: T30608
R;Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.
Science, 273, 813-816, 1996
A;Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host
A;Reference number: Z20876; MUID:96325459
A;Accession: T30608
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1173 <SEN>
A;Cross-references: EMBL:U60315; PIDN:AC55134.1
C;Genetics:
A;Note: MC006L

ALIGNMENTS

C:Genetics:
A:Map position: 7L

Query Match 48.9%; Score 45; DB 2; Length 1403;
Best Local Similarity 63.6%; Pred. No. 25;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 WSAFSCFOKAQ 15
| : ||| || |
Db 361 WTCFSCICLKQK 371

RESULT 3

PSDG
phospholipase A2 (EC 3.1.1.4) precursor - dog
N:Alternate names: phosphatidylcholine 2-acylhydrolase
C:Species: Canis lupus familiaris (dog)
C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jun-2000
C:Accession: S11316; A24392; J50006
R:Kerfelec, B.; LaForge, K.S.; Vasiloudes, P.; Puigserver, A.; Scheele, G.A.
Eur. J. Biochem. 190, 299-304, 1990
A:Title: Isolation and sequence of the canine pancreatic phospholipase A(2) gene.
A:Reference number: S11316; MUID:90306027
A:Accession: S11316
A:Molecule type: DNA
A:Residues: 1-146 <KE2>
R:Kerfelec, B.; LaForge, K.S.; Puigserver, A.; Scheele, G.
Pancreas 1, 430-437, 1986
A:Title: Primary structures of canine pancreatic lipase and phospholipase A2 messenger RNA
A:Reference number: A93751; MUID:87175472
A:Accession: A24392
A:Molecule type: mRNA
A:Residues: 1-146 <KE2>
A:CROSS-references: GB:M35301; NID:g164041; PIDN:AAA30883.1; PID:g164042
R:Ohara, O.; Tamaki, M.; Nakamura, E.; Tsuruta, Y.; Fujii, Y.; Shin, M.; Teraoka, H.; Ok
J. Biochem. 99, 733-739, 1986
A:Title: Dog and rat pancreatic phospholipases A2: complete amino acid sequences deduced
A:Reference number: A92008; MUID:86223862
A:Accession: J50006
A:Molecule type: mRNA
A:Residues: 1-146 <OHA>
A:CROSS-references: GB:D00035; NID:g217659; PIDN:BAA00023.1; PID:g217660
C:Function:

A:Description: catalyzes hydrolysis of 1,2-diacyl-sn-glycero-3-phosphocholine to 1-acyl-
A:Note: the reaction is strongly enhanced when the phospholipid is condensed into a micelle
C:Superfamily: phospholipase A2
C:Keywords: calcium; carboxylic ester hydrolase; lipid degradation; lipid digestion; met
F:1-15/Domain: signal sequence #status predicted <Sig>
F:16-22/Domain: activation peptide #status predicted <APT>
F:23-146/Product: phospholipase A2 #status predicted <MPT>
F:33-99,49-146,51-67,66-127,73-120,83-113,106-118/Disulfide bonds: #status predicted
F:50,52,54,71/Binding site: calcium (Tyr, Gly, Asp) #status predicted
F:70,121/Active site: His, Asp #status predicted

Query Match 45.7%; Score 42; DB 1; Length 146;
Best Local Similarity 61.5%; Pred. No. 9.9;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 NCEWSAFSCFOKA 14
| : ||| || |
Db 119 NCDRSAICFSA 131

RESULT 4

TVCHMC
transforming protein N-myc - chicken
C:Species: Gallus gallus (chicken)
C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jun-2000
C:Accession: A34703; J70545
R:Sawai, S.; Kato, K.; Wakamatsu, Y.; Kondoh, H.
Mol. Cell. Biol. 10, 2017-2026, 1990

A:Title: Organization and expression of the chicken N-myc gene.
A:Reference number: A34703; MUID:90220586

A:Accession: A34703
A:Molecule type: DNA
A:Residues: 1-441 <SAW>
A:CROSS-references: GB:D90071; NID:g222842; PIDN:BAA14112.1; PID:g222843
C:Genetics:
A:Gene: N-myc
A:Introns: 243/1
C:Superfamily: myc transforming protein; myc transforming protein homology
C:Keywords: DNA binding; leucine zipper; nucleus; phosphoprotein; transforming protei
F:11-441/Domain: myc transforming protein homology <MYC>
F:410-438/Region: leucine zipper motif

Query Match 45.7%; Score 42; DB 1; Length 441;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 NCEWSAFSCFOKAQ 15
| : ||| || |
Db 128 DCWWSAFSAREKLE 141

RESULT 5

T51172
betaine-aldehyde dehydrogenase (EC 1.2.1.8) [imported] - Amaranthus hypochondriacus
C:Species: Amaranthus hypochondriacus
C>Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 18-Aug-2000
C:Accession: T51172
R:Legaria, J.; Rajsbbaum, R.; Munoz-Clares, R.A.; Villegas-Sepulveda, N.; Simpson, J.;
Gene 218, 69-76, 1998
A:Title: Molecular characterization of two genes encoding betaine aldehyde dehydrogen
A:Reference number: 225322; MUID:98426159
A:Accession: T51172
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-500 <LEG>
A:CROSS-references: EMBL:AF017150; PIDN:AAB70010.1
A:Experimental source: tissue-type leaf
C:Genetics:
A:Gene: ahybadh4
C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology
C:Keywords: oxidoreductase

Query Match 45.7%; Score 42; DB 2; Length 500;
Best Local Similarity 75.0%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 EWSAFSCF 11
| : ||| || |
Db 280 EWTAFGCF 287

RESULT 6

S57273
lignin peroxidase (EC 1.11.1.-) LP7 precursor - white-rot fungus (Trametes versicolor
C:Species: Trametes versicolor (white-rot fungus)
C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 19-May-2000
C:Accession: S57273
R:Johansson, T.; Nyman, P.O.
Biochim. Biophys. Acta 1263, 71-74, 1995
A:Title: The gene from the white-rot fungus Trametes versicolor encoding the lignin p
A:Reference number: S57273; MUID:95359206
A:Accession: S57273
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-368 <JOH>
A:CROSS-references: EMBL:Z30667; NID:g495282; PIDN:CAA83147.1; PID:g495283
C:Genetics:
A:Introns: 21/1; 90/2; 135/2; 317/3; 342/3; 364/2
C:Superfamily: lignin peroxidase

C; Superfamily: laminin-type EGF-like homology
C; Keywords: alternative splicing; basement membrane; extracellular matrix; glycoprote
F; 1-21/Domain: signal sequence #status predicted <SIG>
F; 22-1193/Product: laminin B2t chain #status predicted <MAT>
F; 84-138/Domain: laminin-type EGF-like homology <LEG1>
F; 517-570/Domain: laminin-type EGF-like homology <LEG>

Query Match 44.68; Score 41; DB 2; Length 1193;
 Best Local Similarity 77.8%; Pred. No. 94;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 NCEWAFSCF 10
 ||| ||||
 Db 601 NCEWAFSCF 609
 ||| ||||
 RESULT 11
 J01899
 genome polyprotein - papaya ringspot virus
 N:Contains: 46k protein; 63k protein; 6K protein; coat protein; cylindrical inclusion pr
 C:Species: papaya ringspot virus, PRSV
 C:Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 17-Nov-2000
 C:Accession: J01899; PQ0509; A56603; S24785; S35725
 R:Yeh, S.D.; Jan, F.J.; Chiang, C.H.; Doong, T.J.; Chen, M.C.; Chung, P.H.; Bau, H.J.
 J. Gen. Virol. 73, 2531-2541, 1992
 A:Title: Complete nucleotide sequence and genetic organization of papaya ringspot virus
 A:Reference number: JQ1899; MUID:93019006
 A:Accession: JQ1899
 A:Molecule type: genomic RNA
 A:Residues: 1-3344 <YEH1>
 A:Cross-references: GB:S46722; NID:g258106; PIDN:AAB23789.1; PID:g258107
 A:Accession: PQ0509
 A:Molecule type: protein
 A:Residues: 548-554;1402-1410 <YEH2>
 R:Arch, C.H.; Yeh, S.D.
 Arch. Virol. 127, 345-354, 1992
 A:Title: Nucleotide sequence comparison of the 3'-terminal regions of severe, mild, and
 A:Reference number: A56603; MUID:93090098
 A:Accession: A56603
 A:Molecule type: genomic RNA
 A:Residues: 2561-3344 <WAN>
 A:Cross-references: EMBL:X67672; NID:g62378; PIDN:CAA47904.1; PID:g62379
 C:Experimental source: type P, strain HA
 C:Superfamily: tobacco etch virus genome polyprotein
 C:Keywords: ATP; coat protein; cylindrical inclusion protein; genome-linked protein; inc
 F:1-547/Product: 63k protein #status predicted <PRO>
 F:548-1004/Product: helper component-proteinase #status predicted <HCP>
 F:1005-1401/Product: 46k protein #status predicted <PRP>
 F:1402-2036/Product: cylindrical inclusion protein #status predicted <CIP>
 F:1486-1493/Region: nucleotide-binding motif A (p-loop)
 F:1571-1576/Region: nucleotide-binding motif B
 F:1575-1578/Region: DEXH motif
 F:2037-2093/Product: 6K protein #status predicted <PRI>
 F:2094-2520/Product: nuclear inclusion a protein #status predicted <NIA>
 F:2521-3037/Product: nuclear inclusion b protein #status predicted <NIB>
 F:3038-3344/Product: coat protein #status predicted <COP>
 F:2156/Binding site: phosphoryl-RNA (Tyr) (covalent) #status predicted
 QY 4 EWSAFSCFQKAL 16
 ||| ||||
 Db 1166 EWRALSFKLHL 1178
 ||| ||||
 RESULT 12
 T17463
 rifamycin polyketide synthase modules 1-3 - Amycolatopsis mediterranei
 C:Species: Amycolatopsis mediterranei
 C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
 C:Accession: T17463
 R:Schupp, T.
 submitted to the EMBL Data Library, December 1997
 A:Reference number: T18802
 A:Accession: T17463

A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-4735 <SCH>
 A:Cross-references: EMBL:AJ223012; NID:e1227119; PID:e1227120; PIDN:CAAL1035.1
 A:Experimental source: strain LBG A3136
 C:Superfamily: acyl carrier protein homology; acetate--CoA ligase homology
 C:Keywords: carrier protein
 F:53-500/Domain: acetate--CoA ligase homology <ACL>
 F:543-610/Domain: acyl carrier protein homology <ACP1>
 F:2102-2173/Domain: acyl carrier protein homology <ACP2>
 F:3079-3150/Domain: acyl carrier protein homology <ACP3>
 F:4578-4649/Domain: acyl carrier protein homology <ACP4>
 Query Match 44.6%; Score 41; DB 2; Length 4735;
 Best Local Similarity 60.0%; Pred. No. 3.2e+02;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 NCEWAFSCF 11
 ||| ||||
 Db 183 NCLWSVASCY 192
 ||| ||||
 RESULT 13
 T48816
 hypothetical protein 15E6.240 [imported] - Neurospora crassa
 C:Species: Neurospora crassa
 C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000
 C:Accession: T48816
 R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: 224541
 A:Accession: T48816
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-142 <SCH>
 A:Cross-references: EMBL:AL353822; GSPDB:GN00112; NCSP:15E6.240
 A:Experimental source: cosmid contig 15E6; strain 74
 C:Genetics:
 A:Gene: NCSP:15E6.240
 A:Map position: 2
 Query Match 43.5%; Score 40; DB 2; Length 142;
 Best Local Similarity 66.7%; Pred. No. 20;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 5 WSAFSCFQK 13
 | |||||
 Db 132 WPFVSCFRK 140
 ||| ||||
 RESULT 14
 T37794
 hypothetical protein SPAC16E8.14c - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
 C:Accession: T37794
 R:Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 submitted to the EMBL Data Library, September 1995
 A:Reference number: 221746
 A:Accession: T37794
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-219 <OLI>
 A:Cross-references: EMBL:298529; PIDN:CAB11042.1; GSPDB:GN00066; SPDB:SPAC16E8.14c
 A:Experimental source: strain 972h-; cosmid c16E8
 C:Genetics:
 A:Gene: SPDB:SPAC16E8.14c
 A:Map position: 1
 A:Introns: 12/2; 32/3; 122/1; 158/3
 C:Superfamily: Neurospora crassa hypothetical protein 13E11.350

Query Match 43.5%; Score 40; DB 2; Length 219;
 Best Local Similarity 56.2%; Pred. NO. 30;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TNCWSAFSCFOKAOL 16
 | | | | | | | | | |
 Db 176 TRCEQSLKSLFKKANL 191

RESULT 15
 T03394
 probable betaine-aldehyde dehydrogenase (EC 1.2.1.8) - rice
 C:Species: Oryza sativa (rice)
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
 C:Accession: T03394
 R:Nakamura, T.; Yokota, S.; Muramoto, Y.; Tsutsui, K.; Oguri, Y.; Fukui, K.; Takabe, T.
 Plant J. 11, 1115-1120, 1997
 A:Title: Expression of a betaine aldehyde dehydrogenase gene in rice, a glycine betaine
 A:Reference number: Z14925; MUID:97336302
 A:Accession: T03394
 A>Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-505 <NAK>
 A:Cross-references: EMBL:AB001348; NID:g2244603; PIDN:BAA21098.1; PID:g2244604
 A:Experimental source: cv. Nipponbare
 C:Genetics:
 A:Introns: 38/1; 85/3; 112/3; 162/2; 193/2; 234/3; 256/3; 279/1; 304/3; 342/3; 368/3; 41
 C:Function:
 A:Description: catalyzes the oxidation of betaine aldehyde to betaine using NAD+ and wat
 A:Pathway: betaine biosynthesis
 A>Note: betaine is a protective osmolyte induced to accumulate under saline or dry condi
 C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology
 C:Keywords: NAD; oxidoreductase; stress-induced protein
 F:48-316/Domain: aldehyde dehydrogenase homology <ALD>

Query Match 43.5%; Score 40; DB 2; Length 505;
 Best Local Similarity 54.5%; Pred. NO. 63;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 EWSAFSCFOKA 14
 | | | | | | | | | |
 Db 282 EWAMFGCFANA 292

Search completed: May 23, 2001, 11:12:53
 Job time: 256 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:20:25 ; Search time 40.06 Seconds
(without alignments)
13.682 Million cell updates/sec

Title: US-09-522-217-2_COPY_69_84
Perfect score: 92
Sequence: 1 TNCWSAFSCFQKAL 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	48.9	1403	1 YGNI_YEAST	P53127 saccharomyc
2	42	45.7	146	1 PA21_CANFA	P06596 canis fami
3	42	45.7	441	1 MYCN_CHICK	P18444 gallus gall
4	42	45.7	501	1 DHAB_AMAHP	O04895 amaranthus
5	41	44.6	516	1 HXK2_DROME	Q9nft7 drosophila
6	41	44.6	1193	1 LMG2_HUMAN	Q13753 homo sapien
7	41	44.6	3344	1 POLG_PRSVH	Q01901 p genome po
8	40	43.5	219	1 YDRE_SCHPO	O13748 schizosacch
9	40	43.5	505	1 DHAB_ORISA	O24174 oryza sativ
10	39	42.4	124	1 PA22_PIG	P04416 sus scrofa
11	39	42.4	132	1 PA21_HORSE	P00594 equus cabal
12	39	42.4	146	1 PA21_CAVPO	P43434 cavia porce
13	39	42.4	146	1 PA21_PIG	P00592 sus scrofa
14	39	42.4	148	1 PA21_HUMAN	P04054 homo sapien
15	39	42.4	223	1 COAT_FRVCA	P05070 tobacco rat
16	39	42.4	1229	1 KPBA_DROME	Q9w391 drosophila
17	39	42.4	1257	1 KPBA_CAEEL	P34335 caenorhabdi
18	39	42.4	1266	1 SYL_HUMAN	P41252 homo sapien
19	38.5	41.8	316	1 YJW3_YEAST	P40587 saccharomyc
20	38	41.3	118	1 PA2A_MICNI	P81166 micrurus ni
21	38	41.3	118	1 PA2B_MICNI	P81167 micrurus ni
22	38	41.3	209	1 HIS2_BACSU	O34912 bacillus su
23	38	41.3	249	1 BA71_BUBSP	P07914 eubacterium
24	38	41.3	427	1 MYCN_SERCA	P26014 serinus can
25	38	41.3	437	1 MYCN_XENLA	P24793 xenopus lae
26	38	41.3	454	1 MYC2_WARMO	P20389 marmota mon
27	38	41.3	454	1 MYC2_SPEBE	Q64210 spermophilu
28	38	41.3	460	1 MYCN_MARMO	Q61976 marmota mon
29	38	41.3	462	1 MYCN_MOUSE	P03966 mus musculu
30	38	41.3	462	1 MYCN_RAT	Q63379 rattus norv
31	38	41.3	464	1 MYCN_HUMAN	P04198 homo sapien
32	38	41.3	470	1 YMK7_YEAST	Q04371 saccharomyc
33	38	41.3	489	1 COAT_FMVD	P09519 figwort mos

34	38	41.3	686	1 FRE1_YEAST	P32791 saccharomyc
35	38	41.3	1228	1 ECM_HUMAN	Q13201 homo sapien
36	38	41.3	2616	1 NDL_DROME	P98159 drosophila
37	37.5	40.8	63	1 NXS2_BUNFA	P14534 bungarus fa
38	37	40.2	346	1 TRA3_LEPBO	O48514 leptospira
39	37	40.2	353	1 FPPS_HUMAN	P14324 homo sapien
40	37	40.2	354	1 KMOS_MSVMT	P32593 moloney mur
41	37	40.2	390	1 KMOS_MOUSE	P00536 mus musculu
42	37	40.2	429	1 MYCS_RAT	P23999 rattus norv
43	37	40.2	497	1 DHAB_SPIOL	P17202 spinacia ol
44	37	40.2	500	1 DHAB_BETVU	P28237 beta vulgar
45	37	40.2	554	1 YDA9_SCHPO	Q10351 schizosacch

ALIGNMENTS

RESULT 1					
YGN1_YEAST	YGN1_YEAST	STANDARD;	PRT;	1403 AA.	
AC	P53127;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	15-DEC-1998 (Rel. 37, Last annotation update)				
DE	HYPOTHETICAL 163.2 KDA PROTEIN IN RPLIB-CEGI INTERGENIC REGION.				
GN	YGL131C OR G2850.				
OS	Saccharomyces cerevisiae (Baker's yeast).				
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.				
OX	NCBI_TaxID=4932;				
RN	[1]				
RP	SEQUENCE OF 1-1232 FROM N.A.				
RC	STRAIN=S288C / FY1679;				
RX	MEDLINE=96437978; PubMed=8840506;				
RA	Escaribano V., Eraso P., Portillo F., Mazon M.J.;				
RT	"Sequence analysis of a 14.6 kb DNA fragment of Saccharomyces cerevisiae chromosome VII reveals SEC27, SSM1b, a putative S-adenosylmethionine-dependent enzyme and six new open reading frames."				
RL	Yeast 12.887-892(1996).				
RN	[2]				
RP	SEQUENCE OF 1388-1403 FROM N.A.				
RC	Cerdan E., Rodriguez-Torres A.M., Rodriguez-Belmonte E., Tizon B.,				
RA	Cadahia J.L., Gonzalez-Siso I.;				
RL	Submitted (MAY-1996) to the EMBL/GenBank/DBAJ databases.				
CC	-!- SIMILARITY: SOME, TO S.POMBE SPAC3H1.12C.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; Z72654; CAA96843.1; -				
DR	EMBL; Z72652; CAA96841.1; -				
DR	SGD; S00030399; YGL131C.				
DR	InterPro; IPR001005; -				
DR	InterPro; IPR001025; -				
DR	InterPro; IPR001965; -				
DR	Pfam; PF01426; BAH; 1.				
DR	Pfam; PF00628; PHD; 2.				
DR	Pfam; PF00249; myb_DNA-binding; 1.				
KW	Hypothetical protein.				
SQ	SEQUENCE 1403 AA; 163202 MW; 758D3FB130DF2357 CRC64;				

Query Match 48.9%; Score 45; DB 1; Length 1403;
Best Local Similarity 63.6%; Pred. No. 9.3;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 5 WSAFSCFQKAL 15

FT QY SEQUENCE 441 AA; 48696 MW; 176B36CE768FC0CE CRC64; (BY SIMILARITY).

Query Match 45.7%; Score 42; DB 1; Length 441;
Best Local Similarity 50.0%; Pred. No. 9.7;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 NCWSAFSFCQKAO 15
: | ||||| : | :
DB 128 DCWSAFSAREKLE 141

RESULT 4
DHAB_AMAHP STANDARD; PRT; 501 AA.
AC O04895;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE BETAINE-ALDEHYDE DEHYDROGENASE PRECURSOR (EC 1.2.1.8) (BADH).
GN BADH4.
OS Amananthus hypochondriacus (Prince's feather).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales;
OC Caryophyllales; Amaranthaceae; Amaranthus.
OX NCBI_TaxID=28502;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=98426159; PubMed=9751804;
RA Legaria J., Rajsbaum R., Munoz-Clares R.A., Villegas-Sepulveda N.,
RA Simpson J., Iturriaga G.;
RT "Molecular characterization of two genes encoding betaine aldehyde
RT dehydrogenase from amaranth. Expression in leaves under short-term
RT exposure to osmotic stress or abscisic acid.";
RL Gene 218:69-76(1998).

CC -!- CATALYTIC ACTIVITY: BETAINE ALDEHYDE + NAD(+) + H(2)O = BETAINE
CC + NADH.
CC -!- PATHWAY: LAST STEP IN THE BIOSYNTHESIS OF BETAINE.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF000132; NAB58165.1; -
CC DR HSP; P20000; I442.
CC InterPro: IPR002086; -
CC Pfam: PF00171; aligned; 1.
CC DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; 1.
CC DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; 1.
CC KW Oxidoreductase; NAD; Chloroplast; Transit peptide.
FT TRANSIT 1 7 CHLOROPLAST (POTENTIAL).
FT CHAIN 8 501 BETAINE-ALDEHYDE DEHYDROGENASE.
FT NP_BIND 238 243 NAD (ADP PART) (BY SIMILARITY).
FT ACT_SITE 260 260 BY SIMILARITY.
FT ACT_SITE 294 294 BY SIMILARITY.
SQ SEQUENCE 501 AA; 54503 MW; D5065A91BFC0D37E CRC64;

Query Match 45.7%; Score 42; DB 1; Length 501;
Best Local Similarity 75.0%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 EWSAFSCF 11
: | : | : | : |

DB 280 EWTAFGCF 287

RESULT 5

HXX2_DROME
ID HXX2_DROME STANDARD; PRT; 516 AA.
AC O9NFT7; O9NFT8; O9VBF1,
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HEXOKINASE TYPE 2 (EC 2.7.1.1).
GN HEX-T2 OR HEX OR CG5443.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VARIOUS STRAINS;
RA Duvernell D.D., Eanes W.F.;
RT "Contrasting molecular population genetics of four hexokinases in
RT Drosophila melanogaster and Drosophila simulans.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-K;
RA Deobagkar D.D., Kulkarni G.V., Deobagkar D.N.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- CATALYTIC ACTIVITY: ATP + D-HEXOSE = ADP + D-HEXOSE 6-PHOSPHATE.
CC -!- PATHWAY: FIRST STEP OF SEVERAL METABOLIC PATHWAYS.

FT	DISULFID	28	37	BY SIMILARITY.
FT	DISULFID	30	53	BY SIMILARITY.
FT	DISULFID	56	65	BY SIMILARITY.
FT	DISULFID	68	81	BY SIMILARITY.
FT	DISULFID	84	96	BY SIMILARITY.
FT	DISULFID	86	102	BY SIMILARITY.
FT	DISULFID	104	113	BY SIMILARITY.
FT	DISULFID	116	128	BY SIMILARITY.
FT	DISULFID	139	150	BY SIMILARITY.
FT	DISULFID	141	155	BY SIMILARITY.
FT	DISULFID	157	166	BY SIMILARITY.
FT	DISULFID	169	184	BY SIMILARITY.
FT	DISULFID	462	470	BY SIMILARITY.
FT	DISULFID	464	481	BY SIMILARITY.
FT	DISULFID	484	493	BY SIMILARITY.
FT	DISULFID	496	514	BY SIMILARITY.
FT	DISULFID	517	531	BY SIMILARITY.
FT	DISULFID	519	538	BY SIMILARITY.
FT	DISULFID	541	550	BY SIMILARITY.
FT	DISULFID	553	570	BY SIMILARITY.
FT	DISULFID	609	609	INTERCHAIN (PROBABLE).
FT	DISULFID	612	612	INTERCHAIN (PROBABLE).
FT	DISULFID	1184	1184	INTERCHAIN (PROBABLE).
FT	CARBOHYD	342	342	N-LINKED (GLCNAC. .)
FT	CARBOHYD	362	362	N-LINKED (GLCNAC. .)
FT	CARBOHYD	942	942	N-LINKED (GLCNAC. .)
FT	CARBOHYD	1033	1033	N-LINKED (GLCNAC. .)
FT	VARSPIC	1110	1193	DPLSVDEGLVLEOKLRKQVINSQLRPMSELEERAR
FT				QQRGHLLHLETSIDGLADVKNLNDRNLPGGCYNTQALE
FT	CONFLICT	12	12	QO -> GM (IN SMALL ISOFORM).
FT	CONFLICT	473	473	F -> L (IN REF. 2).
FT	CONFLICT	521	521	M -> I (IN REF. 3).
FT	CONFLICT	857	857	N -> S (IN REF. 3).
FT	CONFLICT	883	883	R -> P (IN REF. 3).
FT	CONFLICT	883	883	T -> S (IN REF. 2 AND 3).
SQ	SEQUENCE	1193 AA;	130989 MW;	10BFC39E416F4ECE CRC64;

Query Match 44.6%; Score 41; DB 1; Length 1193;

Best Local Similarity 77.8%; Pred. No. 36;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps

Oy	2 NCSEAFSC 10				
	111	1111			
Db	601 NCEHGAFSC 609				

RESULT 7

POLG_PRSVH

ID POLG_PRSVH STANDARD; PRT; 3344 AA.

AC Q01901;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE GENOME POLYPROTEIN [CONTAINS: N-TERMINAL PROTEIN (P1); HELPER

DE COMPONENT PROTEINASE (EC 3.4.22.-) (HC-PRO); PROTEIN P3; 6 KDA PROTEIN

DE 1 (6K1); CYTOPLASMIC INCLUSION PROTEIN (CI); 6 KDA PROTEIN 2 (6K2);

DE GENOME-LINKED PROTEIN (VPG); NUCLEAR INCLUSION PROTEIN A (NI-A) (NIA)

DE (EC 3.4.22.-) (49 KDA PROTEINASE) (49 KDA-PRO); NUCLEAR INCLUSION

DE PROTEIN B (NT-B) (NTB) (RNA-DIRECTED RNA POLYMERASE) (EC 2.7.7.48);

DE COAT PROTEIN (CP);

OS Papaya ringspot virus (strain P / mutant HA)

OS viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;

OC Potyvirus.

OC NCBI_Taxid=31731;

RN [1]

RN SEQUENCE FROM N.A.

RA Wang C.H., Bau H.J., Yeh S.D.;

RT "Comparison of the nuclear inclusion b protein and coat protein genes

RT of five papaya ringspot virus strains distinct in geographic origin

RT and pathogenicity."

RL Phytopathology 84:1205-1210(1994).

RL [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE=93019006; PubMed=1402799;
 RA Yeh S.D., Jan F.J., Chiang C.H., Doong T.J., Chen M.C.,
 RA Chung P.H., Bau H.J.;
 RT "Complete nucleotide sequence and genetic organization of papaya
 RT ringspot virus RNA";
 RL J. Gen. Virol. 73:2531-2541(1992).
 RN [3]
 RP SEQUENCE OF 2561-3344 FROM N.A.
 RX MEDLINE=93090098; PubMed=1456896;
 RA Wang C.H., Yeh S.D.;
 RA "Nucleotide sequence comparison of the 3'-terminal regions of severe,
 RT mild, and non-papaya infecting strains of papaya ringspot virus";
 RL Arch. Virol. 127:345-354(1992).
 RN [3]
 CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID
 CC TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
 CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
 CC MAY BE INVOLVED IN REPLICATION.
 CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
 CC -1- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
 CC -1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
 CC POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
 CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
 CC INDIVIDUAL PROTEINS.
 CC -1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
 CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
 CC -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X67673; CAA47905.1; -;
 DR EMBL; S46722; AAB23789.1; -;
 DR EMBL; X67672; CAA47904.1; -;
 DR PIR; S24785; S24785.
 DR MEROPS; C04.001; -;
 DR MEROPS; C06.001; -;
 DR MEROPS; S30.001; -;
 DR InterPro; IPR001205; -;
 DR InterPro; IPR001456; -;
 DR InterPro; IPR001592; -;
 DR InterPro; IPR001730; -;
 DR InterPro; IPR002540; -;
 DR Pfam; PF00863; Peptidase_C4; 1.
 DR Pfam; PF00851; Peptidase_C6; 1.
 DR Pfam; PF01577; Poty_P1; 1.
 DR Pfam; PF00767; Poty_coat; 1.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR PRINTS; PR00966; NIAPOTYPASE.
 KW Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;
 KW Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
 KW ATP-binding.
 FT CHAIN 1 529 N-TERMINAL PROTEIN (BY SIMILARITY).
 FT CHAIN 530 1149 HELPER COMPONENT PROTEINASE
 FT (BY SIMILARITY).
 FT CHAIN 1150 ? PROTEIN P3 (BY SIMILARITY).
 FT CHAIN ? 1401 6 KDA PROTEIN 1 (BY SIMILARITY).
 FT CHAIN 1402 2036 CYTOPLASMIC INCLUSION PROTEIN (BY
 FT SIMILARITY).
 FT CHAIN 2037 2093 6 KDA PROTEIN 2 (BY SIMILARITY).
 FT CHAIN 2094 ? GENOME-LINKED PROTEIN (BY SIMILARITY).
 FT CHAIN ? 2520 NUCLEAR INCLUSION PROTEIN A
 FT (BY SIMILARITY).
 FT CHAIN 2521 3037 NUCLEAR INCLUSION PROTEIN B
 FT (BY SIMILARITY).
 FT CHAIN 3038 3344 COAT PROTEIN (BY SIMILARITY).
 FT BINDING 2156 COVALENT LINKAGE OF VIRAL RNA (BY
 FT SIMILARITY).

FT NP_BIND 1486 1493 ATP (POTENTIAL).
 SO SEQUENCE 3344 AA; 381040 MW; E90CD7523AC5243D CRC64;
 Query Match 44.6%; Score 41; DB 1; Length 3344;
 Best Local Similarity 61.5%; Pred. No. 93;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 4 EWSAFSCFQKAOL 16
 ||| | | | | |
 Db 1166 EWRALSLFKLHL 1178

 RESULT 8
 YDRE SCHPO
 ID YDRE SCHPO STANDARD; PRT; 219 AA.
 AC O13748;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE HYPOTHETICAL 24.4 KDA PROTEIN C16E8.14C IN CHROMOSOME I.
 GN SPAC16E8.14C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Oliver K., Harris D., Barrell B.G., Rajandream M.A., Wood V.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO YEAST YBR261C.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z98529; CAB11042.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 219 AA; 24388 MW; 3CDEF5386CC90C44 CRC64;
 Query Match 43.5%; Score 40; DB 1; Length 219;
 Best Local Similarity 56.2%; Pred. No. 11;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 Qy 1 TNCWSEAFSCFQKAOL 16
 | | | | | | | | | |
 Db 176 TRCEQSLKSLFKKANL 191

 RESULT 9
 DHAB ORYSA
 ID DHAB ORYSA STANDARD; PRT; 505 AA.
 AC O24174;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE BETAIN-ALDEHYDE DEHYDROGENASE (EC 1.2.1.8) (BADH).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae;
 OC Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RX MEDLINE=97336302; PubMed=9193078;

RA Nakamura T., Yokota S., Muramoto Y., Tsutsui K., Oguri Y., Fukui K.,
 RA Takabe T.;
 RT "Expression of a betaine aldehyde dehydrogenase gene in rice, a
 RT glycinebetaine nonaccumulator, and possible localization of its
 RT protein in peroxisomes.";
 RL Plant J. 11:1115-1120(1997).
 CC -!- CATALYTIC ACTIVITY: BETAINE ALDEHYDE + NAD(+) + H(2)O = BETAINE
 CC + NADH.
 CC
 CC -!- PATHWAY: LAST STEP IN THE BIOSYNTHESIS OF BETAINE.
 CC
 CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC
 CC -!- SUBCELLULAR LOCATION: PEROXISOMAL.
 CC
 CC -!- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AB001348; BAA21098.1; -
 DR InterPro: IPR002086; -
 DR Pfam: pf00171; aldehyd; 1.
 DR PROSITE: PS00342; MICROBODIES_CTER; 1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; 1.
 KW Oxidoreductase; NAD; Chloroplast; Peroxisome.
 FT NP_BIND 240 245 NAD (ADP PART) (BY SIMILARITY).
 FT ACT_SITE 262 262 BY SIMILARITY.
 FT ACT_SITE 296 296 BY SIMILARITY.
 FT SITE 503 505 MICROBODY TARGETING SIGNAL (POTENTIAL).
 FT SEQUENCE 505 AA; 54647 MW; 85EFA42B059A8081 CRC64;
 SQ
 Query Match 43.5%; Score 40; DB 1; Length 505;
 Best Local Similarity 54.5%; Pred. No. 23;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 4 ENSAFSCFOKA 14
 II: | | | |
 DB 282 EWAFCCFANA 292
 RESULT 10
 PA22_PIG STANDARD; PRT; 124 AA.
 ID PA22_PIG
 AC P04416;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PHOSPHOLIPASE A2, MINOR ISOZYME (EC 3.1.1.4) (PHOSPHATIDYLCHOLINE 2-
 DE ACYLHYDROLASE).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP PRELIMINARY SEQUENCE.
 RX MEDLINE=80088382; PubMed=518908;
 RA Puijk W.C., Verheij H.M., Wietzes P., de Haas G.H.;
 RT "The amino acid sequence of the phospholipase A2 isoenzyme from
 RT porcine pancreas";
 RL Biochim. Biophys. Acta 580:411-415(1979).
 CC -!- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE
 CC 2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
 CC -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -!- MISCELLANEOUS: THIS ISOZYME CONSTITUTES ABOUT 5% OF THE PANCREATIC
 CC PHOSPHOLIPASE.
 CC -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
 CC PIR: A00735; PSPGA2.
 DR HSP; P00592; 1SFW.
 DR InterPro: IPR001211; -

DR InterPro: IPR001211; -
 DR Pfam: PF00068; Phoslip; 1.
 DR PRINTS: PR00389; PHPLIPASEA2.
 DR PROSITE: PS00118; PA2_HIS; 1.
 DR PROSITE: PS00119; PA2_ASP; 1.
 KW Hydrolase; Lipid degradation; Pancreas; Calcium.
 FT ACT_SITE 48 48
 FT ACT_SITE 99 99
 FT DISULFID 11 77
 FT DISULFID 27 124
 FT DISULFID 29 45
 FT DISULFID 44 105
 FT DISULFID 51 98
 FT DISULFID 61 91
 FT DISULFID 84 96
 FT CA_BIND 28 28
 FT CA_BIND 30 30
 FT CA_BIND 32 32
 FT CA_BIND 49 49
 FT SEQUENCE 124 AA; 13968 MW; 37E6B121D66FFB2F CRC64;
 SQ
 Query Match 42.4%; Score 39; DB 1; Length 124;
 Best Local Similarity 53.8%; Pred. No. 9.1;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 2 NCEWAFSCFOKA 14
 II: | | | |
 DB 97 NCDNRNAICFSKA 109
 RESULT 11
 PA21_HORSE STANDARD; PRT; 132 AA.
 ID PA21_HORSE
 AC P00594;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PHOSPHOLIPASE A2 PRECURSOR (EC 3.1.1.4) (PHOSPHATIDYLCHOLINE 2-
 DE ACYLHYDROLASE) (GROUP IB PHOSPHOLIPASE A2) (FRAGMENT).
 GN PLA2GIB.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE OF 1-7.
 RX MEDLINE=77134902; PubMed=849461;
 RA Evenberg A., Meyer H., Verheij H.M., de Haas G.H.;
 RT "Isolation and properties of phospholipase A2 and phospholipase A2
 RT from horse pancreas and horse pancreatic juice";
 RL Biochim. Biophys. Acta 491:265-274(1977).
 RN [2]
 RP SEQUENCE OF 8-132, AND ACTIVE SITE.
 RX MEDLINE=77118587; PubMed=838712;
 RA Evenberg A., Meyer H., Gaastria W., Verheij H.M., de Haas G.H.;
 RT "Amino acid sequence of phospholipase A2 from horse pancreas";
 RL J. Biol. Chem. 252:1189-1196(1977).
 RN [3]
 RP REVISIONS TO 131-132.
 RX MEDLINE=83283533; PubMed=6349696;
 RA Verheij H.M., Westerman J., Sternby B., de Haas G.H.;
 RT "The complete primary structure of phospholipase A2 from human
 RT pancreas";
 RL Biochim. Biophys. Acta 747:93-99(1983).
 CC -!- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE
 CC 2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
 CC -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
 CC PIR: A00737; PSHOA.
 DR HSP; P00592; 1SFW.
 DR InterPro: IPR001211; -

Pfam; PF00068; phoslip; 1.
 DR PRINTS: PR00389; PHPLIPASEA2.
 DR PROSITE; PS00118; PA2_HIS; 1.
 DR PROSITE; PS00119; PA2_ASP; 1.
 KW Hydrolase; Lipid degradation; Calcium; Pancreas.
 FT NON_TER 1 7
 FT PROPEP 1 7 REMOVED BY TRYPSIN.
 FT CHAIN 8 132 PHOSPHOLIPASE A2.
 FT ACT_SITE 55 55 BY SIMILARITY.
 FT ACT_SITE 106 106 BY SIMILARITY.
 FT DISULFID 18 84 BY SIMILARITY.
 FT DISULFID 34 131 BY SIMILARITY.
 FT DISULFID 36 52 BY SIMILARITY.
 FT DISULFID 51 112 BY SIMILARITY.
 FT DISULFID 58 105 BY SIMILARITY.
 FT DISULFID 68 98 BY SIMILARITY.
 FT DISULFID 91 103 BY SIMILARITY.
 FT CA_BIND 35 35 HIGH AFFINITY (BY SIMILARITY).
 FT CA_BIND 37 37 HIGH AFFINITY (BY SIMILARITY).
 FT CA_BIND 39 39 HIGH AFFINITY (BY SIMILARITY).
 FT CA_BIND 56 56 HIGH AFFINITY (BY SIMILARITY).
 FT VARIANT 1 2 MISSING (IN A SECOND FORM).
 SQ SEQUENCE 132 AA; 14694 MW; BC6B809A244052EA CRC64;
 Query Match 42.4%; Score 39; DB 1; Length 132;
 Best Local Similarity 53.8%; pred. No. 9.6;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 2 NCSEWAFSCFOKA 14
 Db 104 NCDRNAAICFSKA 116.
 RESULT 12
 ID1_CAVPO
 PA21_CAVPO STANDARD; PRT; 146 AA.
 AC PA4343;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PHOSPHOLIPASE A2 PRECURSOR (EC 3.1.1.4) (PHOSPHATIDYLCHOLINE 2-
 DE ACYLHYDROLASE) (GROUP 1B PHOSPHOLIPASE A2).
 GN PLA2G1B.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 ON NCBI_TaxID=10141;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93345504; PubMed=8344290;
 RA Ying Z., Tojo H., Nonaka Y., Okamoto M.;
 RT Cloning and expression of phospholipase A2 from guinea pig gastric
 RL mucosa, its induction by carbachol and secretion in vivo.";
 RL Eur. J. Biochem. 215:91-97(1993).
 CC -!- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE
 CC 2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
 CC -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
 CC PHOSPHOCHOLINE + A FATTY ACID ANION.
 CC -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. Usage by and for commercial
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; D00740; BRA00640.1; -
 DR HSPSP; P00592; 1SFW.
 DR InterPro; IPR001211; -
 DR Pfam; PF00068; phoslip; 1.

10

FT TURN 139 139
FT HELIX 142 145
SQ SEQUENCE 146 AA; 16279 MW; DE87674C9476FA36 CRC64;

Query Match 42.4%; Score 39; DB 1; Length 146;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 NCWSAFSCFQKA 14
Db 119 NCDRNAAICFCSKA 131
II: : I I I I

RESULT 14
PA21_HUMAN STANDARD; PRT; 148 AA.
AC P04054;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PHOSPHOLIPASE A2 PRECURSOR (BC 3.1.1.4) (PHOSPHATIDYLCHOLINE 2-
DE ACYLHYDROLASE) (GROUP IB PHOSPHOLIPASE A2).
GN PLA2G1B OR PLA2A OR PLA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=87132925; PubMed=3028739;
RA Seilhamer J.J., Randall T.L., Yamanaka M., Johnson L.K.;
RT "Pancreatic phospholipase A2: isolation of the human gene and cDNAs
RT from porcine pancreas and human lung.";
RL DNA 5:519-527(1986).
RN [2]
RP SEQUENCE FROM N.A.
RA Bradshaw H., Wu X., Ozersky P.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 16-22.
RC TISSUE=Pancreas;
RX MEDLINE=82138816; PubMed=7060561;
RA Grataroli R., Dijkman R., Dutilh C.E., van der Ouderaa F.,
RA de Haas G.H., Figarella C.;
RT "Studies on phospholipase A2 and its enzyme from human pancreatic
RT juice, Catalytic properties and sequence of the N-terminal region.";
RL Eur. J. Biochem. 122:111-117(1982).
RN [4]
RP SEQUENCE OF 23-148.
RC TISSUE=Pancreas;
RX MEDLINE=83283533; PubMed=6349696;
RA Verheij H.M., Westerman J., Sternby B., de Haas G.H.;
RT "The complete primary structure of phospholipase A2 from human
RT pancreas.";
RL Biochim. Biophys. Acta 747:93-99(1983).
CC -1- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE
CC 2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
CC PHOSPHOCHOLINE + A FATTY ACID ANION.
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL; M21056; AAA60107.1; -
DR EMBL; M22970; AAA60107.1; JOINED.

DR EMBL; M21054; AAA36450.1; -
DR EMBL; AC003982; AAB95635.1; -
DR PIR; C25793; PSHU.
DR HSSP; P00592; ISFW.
DR MIM; 172410; -
DR InterPro; IPR001211; -
DR Pfam; PF00068; phoslip; 1.
DR PRINTS; PR00389; PHPHLIPASEA2.
DR PROSITE; PS00118; PA2_HIS; 1.
DR PROSITE; PS00119; PA2_HIS; 1.
KW Hydrolase; Lipid degradation; Calcium; Pancreas; Signal.
FT SIGNAL 1 15
FT PROPEP 16 22 ACTIVATION PEPTIDE.
FT CHAIN 23 148 PHOSPHOLIPASE A2.
FT ACT_SITE 70 70 BY SIMILARITY.
FT ACT_SITE 121 121 BY SIMILARITY.
FT DISULFID 33 99
FT DISULFID 49 146
FT DISULFID 51 67
FT DISULFID 66 127
FT DISULFID 73 120
FT DISULFID 83 113
FT DISULFID 106 118
FT CA_BIND 50 50 HIGH AFFINITY (BY SIMILARITY).
FT CA_BIND 52 52 HIGH AFFINITY (BY SIMILARITY).
FT CA_BIND 54 54 HIGH AFFINITY (BY SIMILARITY).
FT CA_BIND 71 71 HIGH AFFINITY (BY SIMILARITY).
FT CONFLICT 144 144 MISSING (IN REF. 4).
SQ SEQUENCE 148 AA; 16360 MW; C8E38B2B64AEE8CB CRC64;

Query Match 42.4%; Score 39; DB 1; Length 148;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 NCWSAFSCFQKA 14
Db 119 NCDRNAAICFCSKA 131
II: : I I I I

RESULT 15
COAT_TRVCA STANDARD; PRT; 223 AA.
AC P05070;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 01-JAN-1988 (Rel. 06, Last annotation update)
DE COAT PROTEIN (CAPSID PROTEIN).
OS Tobacco rattle virus (strain CAM).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobravirus.
OX NCBI_TaxID=12296;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86093657; PubMed=3841203;
RA Bergh S.T., Koziel M.G., Huang S.-C., Thomas R.A., Gilley D.P.,
RA Siegel A.;
RT "The nucleotide sequence of tobacco rattle virus RNA-2 (CAM strain).";
RL Nucleic Acids Res. 13:8507-8518(1985).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL; X03241; CAA26998.1; -
DR PIR; A26027; VCBVCA.
KW Coat protein.
SQ SEQUENCE 223 AA; 23683 MW; EF14652A2F743398 CRC64;

Query Match 42.4%; Score 39; DB 1; Length 223;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 1 TNCWSAFSCFQKAQL 16
 | | | : | | | |
Db 56 TNCPSAEXRFPDKSL 71

Search completed: May 23, 2001, 11:20:26
Job time: 524 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:19:45 ; Search time 118.04 Seconds
(Without alignments)
15.887 Million cell updates/sec

Title: US-09-522-217-2_COPY_69_84
Perfect score: 92
Sequence: 1 TNCWSAFSCFQKAL 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL15.*
1: sp.archaea.*
2: sp.bacteria.*
3: sp.fungi.*
4: sp.human.*
5: sp.invertebrate.*
6: sp.mammal.*
7: sp.mmc.*
8: sp.organelle.*
9: sp.phage.*
10: sp.plant.*
11: sp.rodent.*
12: sp.unclassified.*
13: sp.vertebrate.*
14: sp.virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	47	51.1	226	5 Q9VU26	Q9vU26 drosophila
2	46	50.0	1173	14 Q98177	Q98177 molluscum c
3	46	50.0	1175	14 Q85285	Q85285 molluscum c
4	44	47.8	772	10 Q9SY24	Q9SY24 arabidopsis
5	42	45.7	500	10 Q22477	Q22477 amaranthus
6	42	45.7	738	4 Q9Y3M9	Q9y3M9 homo sapien
7	41.5	45.1	368	3 Q99057	Q99057 trametetes ve
8	41	44.6	254	5 Q9NFT7	Q9nft7 drosophila
9	41	44.6	315	4 Q9UF97	Q9uf97 homo sapien
10	41	44.6	347	4 Q9NW93	Q9nw93 homo sapien
11	41	44.6	354	11 Q82509	Q62509 mus musculus
12	41	44.6	455	11 Q62511	Q62511 mus musculus
13	41	44.6	474	4 Q9NWY6	Q9nwY6 homo sapien
14	41	44.6	530	11 P70196	P70196 mus musculus
15	41	44.6	544	5 Q26858	Q26858 trypanosoma
16	41	44.6	707	4 Q9UJC8	Q9uJC8 homo sapien
17	41	44.6	907	5 Q9VBF1	Q9vbf1 drosophila
18	41	44.6	1002	14 Q85025	Q85025 papaya ring
19	41	44.6	1072	4 Q9Y4G7	Q9y4G7 homo sapien

20	41	44.6	1201	5 Q9WLJ0	Q9wLj0 drosophila
21	41	44.6	1935	5 Q9VQW0	Q9vQw0 drosophila
22	41	44.6	3344	14 P90213	P90213 papaya ring
23	41	44.6	4735	2 Q54666	Q54666 amycolatops
24	40	43.5	142	3 Q9P6S9	Q9p6S9 neurospora
25	40	43.5	258	14 Q01999	Q01999 choristoneu
26	40	43.5	262	14 Q9PX90	Q9px90 rabbit fibr
27	40	43.5	474	13 Q73906	Q73906 gallus gall
28	40	43.5	494	13 Q73621	Q73621 xenopus lae
29	40	43.5	499	5 Q9W282	Q9w282 drosophila
30	40	43.5	533	5 Q9TYL4	Q9tYl4 caenorhabdi
31	40	43.5	677	10 Q9LWM3	Q9lWm3 oryza sativ
32	39	42.4	236	11 Q9Z1Y0	Q9z1Y0 rattus norv
33	39	42.4	292	4 Q9NX52	Q9nx52 homo sapien
34	39	42.4	308	14 Q9JGV0	Q9jgv0 plutella xy
35	39	42.4	332	10 Q9LMH3	Q9lMh3 arabidopsis
36	39	42.4	806	10 Q64752	Q64752 arabidopsis
37	39	42.4	836	10 Q9SAK0	Q9sAk0 arabidopsis
38	39	42.4	1140	10 Q9LR04	Q9lR04 arabidopsis
39	39	42.4	1146	3 Q00048	Q00048 ajeilomyces
40	39	42.4	1457	10 Q9LGF5	Q9lGf5 oryza sativ
41	39	42.4	2395	5 Q27167	Q27167 paramescium
42	39	42.4	4599	11 Q9J118	Q9j118 mus musculu
43	38.5	41.8	1878	5 Q9VD86	Q9vd86 drosophila
44	38	41.3	142	5 Q9VDL0	Q9vDl0 drosophila
45	38	41.3	263	14 Q85296	Q85296 myxoma viru

ALIGNMENTS

RESULT 1

ID	Q9VU26	PRELIMINARY;	PRT;	226 AA.
AC	Q9VU26;			
DT	01-MAY-2000	(TREMBLrel. 13, Created)		
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)		
DE	CG10752	PROTEIN.		
GN	CG10752.			
OS	Drosophila melanogaster	(Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; brachycera; Muscomorpha;			
OC	Ephyroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN=BERKELEY;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,			

```

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reiner K.C., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "the genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003539; AAF49865.1; -
DR FLYBASE; FBgn0036325; CG10752.
SQ SEQUENCE 226 AA; 26205 MW; BDE9D172B1EBD230 CRC64;

Query Match 51.1%; Score 47; DB 5; Length 226;
Best Local Similarity 53.3%; Pred. No. 1.2;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TNCWSAFSCFOKAO 15
|:|:| |:|:| |:|:|
Db 159 THCELSFSCYLKAK 173

RESULT 2
Q98177 PRELIMINARY; PRT; 1173 AA.
AC Q98177;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE M0006L.
GN M0006L.
OS Molluscum contagiosum virus subtype 1 (MCVI).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Molluscipoxvirus.
OX NCBI_TaxID=10280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96325459; PubMed=8670425;
RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,
RA Moss B.;
RT "Genome sequence of a human tumorigenic poxvirus: prediction of
RT specific host response-evasion genes.";
RL Science 273:813-816(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,
RA Moss B.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U60315; AAC55134.1; -
SQ SEQUENCE 1173 AA; 125741 MW; 8C2E487E0B574476 CRC64;

Query Match 50.0%; Score 46; DB 14; Length 1173;
Best Local Similarity 60.0%; Pred. No. 8.1;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 CEWSAFSCFO 12
|:|:| |:|:| |:|:|
Db 803 CEWATFLCFR 812

RESULT 3
Q85285 PRELIMINARY; PRT; 1175 AA.
ID Q85285
AC Q85285;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

```

```

DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE ORF17.
OS Molluscum contagiosum virus subtype 1 (MCVI).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Molluscipoxvirus.
OX NCBI_TaxID=10280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94041989; PubMed=8225909;
RA Hadasch R.P., Bugert J.J., Janssen W., Darai G.;
RT "Characterization of the genome of molluscum contagiosum virus type 1
RT between the genome coordinates 0.045 and 0.075 by DNA nucleotide
RT sequence analysis of a 5.6-kb HindIII/MluI DNA fragment.";
RL Intervirology 36:32-43(1993).
DR EMBL; L10127; AAA16882.1; -
SQ SEQUENCE 1175 AA; 126224 MW; 97D75229C41A1A5E CRC64;

Query Match 50.0%; Score 46; DB 14; Length 1175;
Best Local Similarity 60.0%; Pred. No. 8.1;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 CEWSAFSCFO 12
|:|:| |:|:| |:|:|
Db 806 CEWATFLCFR 815

RESULT 4
Q9SY24 PRELIMINARY; PRT; 772 AA.
ID Q9SY24;
AC Q9SY24;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE T17H7.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Buehler E., Shinn P., Dewar K., Feng J., Kim C., Li Y., Sun H.,
RA Conway A., Conway A., Kurtz D., Oji O., Shen Y.K., Toriumi M.,
RA Vysotskaia V., Yu G., Davis R.W., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC T17H7 from Chromosome
RT 1.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004135; AAD32935.1; -
SQ SEQUENCE 772 AA; 87297 MW; 503781B87C1A090B CRC64;

Query Match 47.8%; Score 44; DB 10; Length 772;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TNCWSAFSCFOKAO 16
|:|:| |:|:| |:|:|
Db 466 SNCERECFCFYDLHL 481.

RESULT 5
Q22477 PRELIMINARY; PRT; 500 AA.
ID Q22477;
AC Q22477;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE BETAINA ALDEHYDE DEHYDROGENASE (EC 1.2.1.8).
GN AHYBADH4.
OS Amananthus hypochondriacus (Prince's feather).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

```

OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
 OC Caryophyllales; Amaranthaceae; Amaranthus.
 OX NCBI_TaxID=28502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LEAF;
 RA Legaria J., Iturriaga G.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF017150; AAB70010.1; -;
 DR HSSP; P20000; 1AG8.
 DR MENDEL; 26369; Amahp;1356;26369.
 DR INTERPRO; IPR002086; -;
 DR PFAM; PF00171; aldedb; 1.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 500 AA; 54435 MW; CE7F70B1B288DCAB CRC64;

Query Match 45.7%; Score 42; DB 10; Length 500;
 Best Local Similarity 75.0%; Pred. No. 18;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 EWSAFSCF 11
 ||:||||
 Db 280 EWTAFGCF 287

RESULT 6
 O9Y3M9 PRELIMINARY; PRT; 738 AA.
 ID O9Y3M9;
 AC O9Y3M9;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE DJ694B14.1 (PUTATIVE NOVEL KRAB BOX PROTEIN WITH 18 C2H2 TYPE ZINC
 DE FINGER DOMAINS) (FRAGMENT).
 GN DJ694B14.1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith M.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL031673; CAB43548.1; -;
 DR HSSP; P08045; 1ZNF.
 DR INTERPRO; IPR000822; -;
 DR INTERPRO; IPR001909; -;
 DR PFAM; PF00096; zf-C2H2; 20.
 DR PFAM; PF01352; KRAB; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2; 19.
 KW Zinc-finger; Metal-binding; DNA-binding.
 FT NON_TER 1
 FT NON_TER 738 738
 SQ SEQUENCE 738 AA; 85532 MW; B8443230C27EPD39 CRC64;

Query Match 45.7%; Score 42; DB 4; Length 738;
 Best Local Similarity 42.9%; Pred. No. 26;
 Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 NCWSAFSCFQKAO 15
 | | | | | : |
 Db 166 NSRWGAFKCAERQ 179

RESULT 7
 O99057 PRELIMINARY; PRT; 368 AA.
 ID O99057;
 AC O99057;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE LIGNIN PEROXIDASE ISOZYME LP7 PRECURSOR.
 GN LPGVI.
 OS Trametes versicolor (White-rot fungus).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Aphyllophorales;
 OC Coriolaceae; Trametes.
 OX NCBI_TaxID=5325;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PRL 572; TISSUE=MYCELIUM;
 RX MEDLINE=95359206; PubMed=7632736;
 RA Johansson T., Nyman P.;
 RL "The gene from the white-rot fungus Trametes versicolor encoding the
 RL lignin peroxidase isozyme LP7.";
 RL Biochim. Biophys. Acta 1263:71-74(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PRL 572; TISSUE=MYCELIUM;
 RX MEDLINE=93143365; PubMed=8424691;
 RA Johansson T., Wellinder K., Nyman P.;
 RL "Isozymes of lignin peroxidase and manganese(II) peroxidase from the
 RL white-rot basidiomycete Trametes versicolor. II. Partial sequences,
 RL peptide maps, and amino acid and carbohydrate compositions.";
 RL Arch. Biochem. Biophys. 300:57-62(1993).
 DR EMBL; Z30667; CAA83147.1; -;
 DR HSSP; P06181; 1B82.
 DR INTERPRO; IPR001621; -;
 DR INTERPRO; IPR002016; -;
 DR PFAM; PF00141; peroxidase; 1.
 DR PRINTS; PR00458; PEROXIDASE.
 DR PRINTS; PR00462; LIGNINASE.
 DR PROSITE; PS00436; PEROXIDASE_2; 1.
 KW Signal; peroxidase. 26
 FT SIGNAL 1
 FT CHAIN 27 368 LIGNIN PEROXIDASE ISOZYME LP7.
 SQ SEQUENCE 368 AA; 39319 MW; 4757EE49D97C112C CRC64;

Query Match 45.1%; Score 41.5; DB 3; Length 368;
 Best Local Similarity 56.2%; Pred. No. 17;
 Matches 9; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 TNCEWSAF-SCFQKAO 15
 | : | | | : | | | |
 Db 276 TSCEWQSFVNQKAO 291

RESULT 8
 O9NFT7 PRELIMINARY; PRT; 254 AA.
 ID O9NFT7
 AC O9NFT7
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE HEXOKINASE (EC 2.7.1.1).
 GN HEX.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OREGON K;
 RA Deobackar D.D., Kulkarni G.V., Deobackar D.N.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ271350; CAB72132.1; -;
 KW Kinase; Transferase.
 SQ SEQUENCE 254 AA; 28014 MW; 4DEBD9A5D84FB277 CRC64;

Query Match 44.6%; Score 41; DB 5; Length 254;

Best Local Similarity 85.7%; Pred. No. 15; Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NCWSAF 8
|||||
Db 48 NCWGA 54

RESULT 9

Q9UF97 Q9UF97 PRELIMINARY; PRT; 315 AA.
AC Q9UF97;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE HYPOTHETICAL 34.4 KDA PROTEIN.
GN DKFZP434G1221.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL133100; CAB61410.1; -;
DR HSSP; P08046; IAI1.
DR INTERPRO; IPR000822; -;
DR PFAM; PF00096; zf-C2H2; 5.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 5.
KW Hypothetical protein.
SQ SEQUENCE 315 AA; 34380 MW; 8A85599F47CE19BA CRC64;

Query Match 44.6%; Score 41; DB 4; Length 315;
Best Local Similarity 57.1%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CEWSAFSCFQKQAL 16
||| |||||
Db 215 CEICGFTCRQKASL 228

RESULT 10

Q9NW93 Q9NW93 PRELIMINARY; PRT; 347 AA.
AC Q9NW93;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE CDNA FLJ10208 FIS, CLONE HEMBA1006248, WEAKLY SIMILAR TO ZINC FINGER PROTEIN MFG-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=WHOLE EMBRYO, MAINLY HEAD;
RA Isoyaki T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y., Oshima A.;
RT "NEDO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK001070; BAA91490.1; -;
SQ SEQUENCE 347 AA; 38690 MW; 2E183849B1FFE093 CRC64;

Query Match 44.6%; Score 41; DB 4; Length 347;
Best Local Similarity 57.1%; Pred. No. 20;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CEWSAFSCFQKQAL 16
||| |||||
Db 247 CEICGFTCRQKASL 260

RESULT 11

Q62509 Q62509 PRELIMINARY; PRT; 354 AA.
AC Q62509;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE ZINC FINGER PROTEIN (FRAGMENT).
GN PZF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA MEDLINE-95137394; PubMed-7835706;
RA Saotome Y., Winter C.G., Hirsh D.;
RT "A widely expressed novel C2H2 zinc-finger protein with multiple consensus phosphorylation sites is conserved in mouse and man."
RL Gene 152:233-238(1995).
DR EMBL; U05342; AAA81911.1; -;
DR HSSP; P03001; ITF3.
DR INTERPRO; IPR000822; -;
DR PFAM; PF00096; zf-C2H2; 5.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 5.
KW Zinc-finger; Metal-binding; DNA-binding.
FT NON_TER 1
SQ SEQUENCE 354 AA; 40636 MW; 77AFB496902328D4 CRC64;

Query Match 44.6%; Score 41; DB 11; Length 354;
Best Local Similarity 57.1%; Pred. No. 20;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CEWSAFSCFQKQAL 16
||| |||||
Db 229 CEICGFTCRQKASL 242

RESULT 12

Q62511 Q62511 PRELIMINARY; PRT; 455 AA.
AC Q62511;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE ZINC FINGER PROTEIN PZF.
GN PZF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE=BRAIN;
RA MEDLINE-95137394; PubMed-7835706;
RA Saotome Y., Winter C.G., Hirsh D.;
RT "A widely expressed novel C2H2 zinc-finger protein with multiple consensus phosphorylation sites is conserved in mouse and man."
RL Gene 152:233-238(1995).
DR EMBL; U05343; AAA81913.1; -;
DR HSSP; P03001; ITF3.
DR INTERPRO; IPR000822; -;
DR PFAM; PF00096; zf-C2H2; 5.
DR PRINTS; PR00048; ZINC_FINGER.

DR PROSITE; PS00028; ZINC_FINGER_C2H2; 5.
 KW Zinc-finger; Metal-binding; DNA-binding.
 SQ SEQUENCE 455 AA; 51540 MW; F27A84F2937E223C CRC64;

Query Match 44.6%; Score 41; DB 11; Length 455;
 Best Local Similarity 57.1%; Pred. No. 25;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 CEWSAFSCFQKQAL 16
 || | | | | | |
 Db 287 CEICGTCRQKASL 300

RESULT 13

Q9NMY6 PRELIMINARY; PRT; 474 AA.
 AC Q9NMY6;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE CDNA FLJ20531 FIS, CLONE KAT10870.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
 RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Isoigai T., Sugano S.;
 RT "NEO human cDNA sequencing project."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AK000538; BAA91239.1;
 SQ SEQUENCE 474 AA; 51904 MW; 53B01E31B3E23AFD CRC64;

Query Match 44.6%; Score 41; DB 4; Length 474;
 Best Local Similarity 57.1%; Pred. No. 26;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 CEWSAFSCFQKQAL 16
 || | | | | | |
 Db 374 CEICGTCRQKASL 387

RESULT 14

P70196 PRELIMINARY; PRT; 530 AA.
 AC P70196;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE TRAF6.
 GN TRAF6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Inoue J.;
 RT "TRAF6, a novel CD40 associated factor.";
 RL J. Biol. Chem. 0:0-0(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97067112; PubMed=8910514;
 RA Ishida T., Mizushima S., Azuma S., Kobayashi N., Tojo T., Suzuki K.,
 RA Aizawa S., Watanabe T., Mosialos G., Kieff E., Yamamoto T., Inoue J.;
 RT "Identification of TRAF6, a novel tumor necrosis factor receptor-
 RT associated factor protein that mediates signaling from an amino-
 RT terminal domain of the CD40 cytoplasmic region.";
 RL J. Biol. Chem. 271:28745-28748(1996).

CC -1- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
 DR EMBL; D84655; BAA12705.1;
 DR MGD; MGI:108072; Traf6.
 DR INTERPRO; IPR000504;
 DR INTERPRO; IPR001293;
 DR INTERPRO; IPR001841;
 DR INTERPRO; IPR002083;
 DR INTERPRO; IPR003007;
 DR PFAM; PF00097; zf-C3HC4; 1.
 DR PFAM; PF00917; MATH; 1.
 DR PFAM; PF02176; zf-TRAF; 2.
 DR PROSITE; PS00030; RNP.1; UNKNOWN.1.
 DR PROSITE; PS00518; ZINC_FINGER_C3HC4; 1.
 KW Zinc-finger.
 SQ SEQUENCE 530 AA; 60082 MW; 092D820B4CEDB85B CRC64;

Query Match 44.6%; Score 41; DB 11; Length 530;
 Best Local Similarity 53.8%; Pred. No. 29;
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 CEWSAFSCFQKQAL 15
 | | | | | | |
 Db 242 CTFSVFGCHQKMQ 254

RESULT 15

Q26858 PRELIMINARY; PRT; 544 AA.
 AC Q26858;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE HEXOSE TRANSPORTER.
 OS Trypanosoma cruzi.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5693;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CL;
 RX MEDLINE=94336729; PubMed=8058795;
 RA Tetaud E., Bringaud F., Chabas S., Barret M., Baltz T.;
 RT "Characterization of glucose transport and cloning of a hexose
 RT transporter gene in Trypanosoma cruzi.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:8278-8282(1994).
 DR EMBL; U05588; AAA21207.1;
 DR INTERPRO; IPR000504;
 DR INTERPRO; IPR001066;
 DR PFAM; PF00083; sugar_tr; 1.
 DR PRINTS; PR00171; SUGRTRNSPORT.
 DR PROSITE; PS00030; RNP.1; UNKNOWN.1.
 SQ SEQUENCE 544 AA; 58733 MW; 395F69DC3DD0E8A9 CRC64;

Query Match 44.6%; Score 41; DB 5; Length 544;
 Best Local Similarity 55.6%; Pred. No. 29;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 CEWSAFSCF 11
 | | | | | | |
 Db 87 CGWNGFTCF 95

Search completed: May 23, 2001, 11:19:46
 Job time: 599 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:11:37 ; Search time 108.07 Seconds
(without alignments)
7.405 Million cell updates/sec

Title: US-09-522-217-2_COPY_92_105
Perfect score: 66
Sequence: 1 NERIINVSIIKKLR 14

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	66	100.0	162 21 B18623	A human zalphall 1
2	66	100.0	519 21 B18627	Amino acid sequenc
3	41	62.1	486 21 G53097	Arabidopsis thalia
4	41	62.1	584 21 G53096	Arabidopsis thalia
5	41	62.1	615 21 G53095	Arabidopsis thalia
6	38	57.6	1245 21 B18244	Plasmodium falcipa
7	37	56.1	382 17 R90551	pJG4-5-CDK-BP clon
8	37	56.1	408 20 Y42419	Jun N-terminal kin
9	37	56.1	423 17 R89408	Stress activated p
10	37	56.1	423 17 R89409	Stress activated p
11	37	56.1	443 20 Y34467	Porphyromonas ging

12	37	56.1	451	20	Y34343
13	37	56.1	540	16	R76062
14	37	776	15	15	R60178
15	37	56.1	787	16	R76061
16	36	54.5	232	21	B52547
17	36	54.5	1123	20	Y41086
18	36	54.5	1123	21	B15724
19	35	53.0	218	20	W89435
20	35	53.0	218	20	W83371
21	35	53.0	218	21	Y57685
22	35	53.0	224	20	Y05738
23	35	53.0	233	18	W26639
24	35	53.0	270	18	W06484
25	35	53.0	383	21	Y90863
26	35	53.0	411	14	R40846
27	35	53.0	508	21	B18187
28	35	53.0	512	21	B24094
29	35	53.0	535	21	B59021
30	35	53.0	2308	15	R52580
31	35	53.0	2308	15	R57902
32	34	51.5	102	21	Y93063
33	34	51.5	128	21	Y93238
34	34	51.5	139	21	B07790
35	34	51.5	139	21	B07800
36	34	51.5	211	10	P90412
37	34	51.5	212	20	W96293
38	34	51.5	384	16	R70767
39	34	51.5	384	17	R89411
40	34	51.5	384	20	Y42420
41	34	51.5	384	20	Y42422
42	34	51.5	384	21	B07788
43	34	51.5	384	21	B07798
44	34	51.5	384	21	Y94502
45	34	51.5	388	21	B56931

ALIGNMENTS

RESULT 1
B18623
ID B18623 standard; Protein; 162 AA.
XX
AC B18623;
XX
DT 22-JAN-2001 (first entry)
XX
DE A human zalphall ligand polypeptide.
XX
KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
KW tumourigenesis; leukaemia; hematopoesis; B cell tumour.
XX
OS Homo sapiens.
XX
PN WO200053761-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-US06067.
XX
PR 09-MAR-1999; 99US-0264908.
PR 11-MAR-1999; 99US-0265992.
PR 01-JUL-1999; 99US-0142013.
XX
(ZYMO) ZYMOGENETICS INC.
XX
PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Gross JA;
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX
DR WFI; 2000-565600/52.
DR N-PSDB; A75552.
XX
PT New human cytokine, designated zalphall ligand, useful for stimulating

PT the proliferation and/or development of haematopoietic cells in vitro
PT and in vivo, and for treating tumourigenesis -
FS Disclosure: Page 205-206; 256pp; English.
XX
CC The present sequence represents a human zalphall ligand polypeptide,
CC which is a cytokine. The zalphall ligand is useful for stimulating the
CC proliferation and development of haematopoietic cells in vitro and in
CC vivo. Zalphall ligand polynucleotides can be used as primers or probes
CC for cloning the zalphall gene. The zalphall ligand is useful for
CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
CC used for treating leukaemias and lymphomas. Antagonists against zalphall
CC ligand are useful as research reagents for characterizing ligand-receptor
CC interaction. Antagonists are also useful for inhibiting expansion,
CC proliferation, activation and differentiation of cells involved in
CC regulating hematopoiesis. The zalphall ligand may also be used to
CC stimulate an immune response against B cell tumour, a virus, a parasite
CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
CC agonists and antibodies are also useful for the detection, diagnosis,
CC prevention, and treatment of diseases associated with a zalphall ligand
CC genetic defect.
XX
SQ Sequence 162 AA;

Query Match 100.0%; Score 66; DB 21; Length 162;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NERIINVSIIKKLR 14
Db 92 neriinvsikkllkr 105
|||||

RESULT 2
B18627
ID B18627 standard; Protein; 519 AA.
XX AC B18627;
XX DT 22-JAN-2001 (first entry)
XX DE Amino acid sequence of MBP-human zalphall ligand fusion in pTAP126.
XX KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
XX KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
XX OS Synthetic.
XX OS Homo sapiens.
XX FN WO200053761-A2.
XX PD 14-SEP-2000.
XX PF 09-MAR-2000; 2000WO-US06067.
XX PR 09-MAR-1999; 99US-0264908.
XX PR 11-MAR-1999; 99US-0265992.
XX PR 01-JUL-1999; 99US-0142013.
XX PA (ZYMO) ZYMOGENETICS INC.
XX PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
XX PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX PI WPI: 2000-565600/52.
XX DR N-PSDB; A75599.
XX XX New human cytokine, designated zalphall ligand, useful for stimulating
XX PT the proliferation and/or development of haematopoietic cells in vitro
XX PT and in vivo, and for treating tumourigenesis -
XX XX Example 31; Page 233-235; 256pp; English.

XX The present sequence represents a MBP-human zalphall ligand fusion in
CC the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
CC is useful for stimulating the proliferation and development of
CC haematopoietic cells in vitro and in vivo. Zalphall ligand
CC polynucleotides can be used as primers or probes for cloning the
CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
CC A zalphall ligand-saporin fusion toxin may be used for treating
CC leukaemias and lymphomas. Antagonists against zalphall ligand are
CC useful as research reagents for characterizing ligand-receptor
CC interaction. Antagonists are also useful for inhibiting expansion,
CC proliferation, activation and differentiation of cells involved in
CC regulating hematopoiesis. The zalphall ligand may also be used to
CC stimulate an immune response against B cell tumour, a virus, a parasite
CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
CC agonists and antibodies are also useful for the detection, diagnosis,
CC prevention, and treatment of diseases associated with a zalphall ligand
CC genetic defect.
XX
SQ Sequence 519 AA;

Query Match 100.0%; Score 66; DB 21; Length 519;
Best Local Similarity 100.0%; Pred. No. 0.00076;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NERIINVSIIKKLR 14
Db 449 neriinvsikkllkr 462
|||||

RESULT 3
G53097
ID G53097 standard; Protein; 486 AA.
XX AC G53097;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 67568.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 28-APR-1999; 99US-0130891.
XX PR 30-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 04-MAY-1999; 99US-0132407.
XX PR 05-MAY-1999; 99US-0132484.
XX PR 06-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144532.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 26-JUL-1999; 99US-0145224.
 PR 27-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.

PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match 62.1%; Score 41; DB 21; Length 584;		
Best Local Similarity 57.1%; Pred. No. 23;		
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;		
QY	1 NERIINWSIKLKR 14	
Db	319 neavandniklkk 332	
: : :		
RESULT 5		
G53095	G53095 standard; Protein; 615 AA.	
XX	G53095;	
AC	G53095;	
XX	18-OCT-2000 (first entry)	
DT	Arabidopsis thaliana protein fragment SEQ ID NO: 67566.	
XX	Arabidopsis thaliana.	
DE	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX	Arabidopsis thaliana.	
OS	EP1033405-A2.	
XX	06-SEP-2000.	
PN	25-FEB-2000; 2000EP-0301439.	
PD	25-FEB-1999;	
PF	99US-0121825.	
XX	99US-0123180.	
PR	99US-0123548.	
PR	99US-0125788.	
PR	99US-0126264.	
PR	99US-0126785.	
PR	99US-0127462.	
PR	99US-0128234.	
PR	99US-0128714.	
PR	99US-0129845.	
PR	99US-0130077.	
PR	99US-0130449.	
PR	99US-0130510.	
PR	99US-0130891.	
PR	99US-0131449.	
PR	99US-0132048.	
PR	99US-0132407.	
PR	99US-0132484.	
PR	99US-0132485.	
PR	99US-0132486.	
PR	99US-0132487.	
PR	99US-0132863.	
PR	99US-0134256.	
PR	99US-0134218.	
PR	99US-0134219.	
PR	99US-0134221.	
PR	99US-0134370.	
PR	99US-0134768.	
PR	99US-0134941.	
PR	99US-0135124.	
PR	99US-0135353.	
PR	99US-0135629.	
PR	99US-0136021.	
PR	99US-0136392.	

PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145931.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 57.6%; Score 41; DB 21; Length 615;
 Best Local Similarity 57.1%; Pred. No. 24;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NERIINSIKKLR 14
 || : |||||
 Db 350 neavandhkkkk 363

RESULT 6
 ID B18244 standard; Protein; 1245 AA.
 XX
 AC B18244;
 XX
 DT 07-NOV-2000 (first entry)
 XX
 DE Plasmodium falciparum chromosome 2 related protein SEQ ID NO:101.
 XX
 KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
 KW antimalarial; malaria; protozoacide; infection; insecticide.
 XX
 OS Plasmodium falciparum.
 XX
 PN WO200025728-A2.
 XX
 PD 11-MAY-2000.
 XX
 PF 05-NOV-1999; 99WO-US26796.
 XX
 PR 05-NOV-1998; 98US-0107131.
 XX
 PA (HOFF/) HOFFMAN S.
 PA (CARU/) CARUCCI D.
 PA (GARD/) GARDNER M.
 PA (VENT/) VENTER J C.
 XX
 PI Hoffman S, Carucci D, Gardner M, Venter JC;
 XX
 DR WPI; 2000-365347/31.
 XX
 PT Proteins encoded by chromosome 2 of the human malarial parasite,
 PT Plasmodium falciparum, useful as antimalarial vaccines and in the
 PT diagnosis of P.falciparum infection -
 XX
 PS Disclosure; Page 234-237; 577pp; English.
 XX
 CC The present invention describes proteins and their fragments (I) encoded
 CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
 CC Also described are: (I) nucleotide sequences (II) encoding (I); and (2)
 CC vaccines against P. falciparum infection comprising (I) or (II).
 CC (I) and (II) are useful for the development of vaccines against
 CC P. falciparum infection. (I) and polyclonal antisera or a monoclonal
 CC antibody raised to immunogens comprising the sequences of (I), are
 CC useful in the detection of infection with P. falciparum. Furthermore,
 CC (I) (especially when they are rifins or secreted or membrane proteins)
 CC can aid the identification of drugs to treat or prevent P. falciparum
 CC infection, or they can be used to identify drug resistance in
 CC P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
 CC subsequent identification of proteins encoded by it will help to expand
 CC our understanding of parasite biology, a process hampered by the
 CC complexity of the parasitic lifecycle, and provide new targets for
 CC vaccine and drug development. Parasite resistance to drugs and mosquito
 CC resistance to insecticides have led to a resurgence of malaria in many
 CC parts of the world, and there is a pressing need for vaccines and new
 CC drugs. A70078 to A70287 and B18144 to B18352 represent nucleotide and
 CC protein sequences given in the present invention, but which are not
 CC specifically mentioned within the specification.
 XX
 SQ Sequence 1245 AA;

Query Match 57.6%; Score 38; DB 21; Length 1245;
 Best Local Similarity 60.0%; Pred. No. 1.7e+02;
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 NERIINSIK 10
 ||| : ||| : |
 Db 2 nervlnvank 11

RESULT 7
 ID R90551 standard; protein; 382 AA.
 XX
 AC R90551;
 XX
 DT 08-AUG-1996 (first entry)
 XX
 DE pJG4-5-CDK-BP clone #227 derived CDK4 binding protein.
 XX
 KW Cell cycle; CDK4; regulation; G1 phase; proliferation; tumourigenesis;
 KW cyclin dependent kinase; differentiation; CDK4 inhibitor; agonist;
 KW antagonist.
 XX
 OS Synthetic.
 XX
 PN WO9533819-A2.
 XX
 PD 14-DEC-1995.
 XX
 PF 02-JUN-1995; 95WO-US07113.
 XX
 PR 02-JUN-1994; 94US-0253155.
 XX
 PA (MITO-) MITOTIX INC.
 XX
 XX Draetta G, Gyuris J;
 PI WPI; 1996-040227/04.
 DR N-PSDB; T12180.
 XX
 PT Cyclin-dependent kinase-4 binding protein - used in the isolation of
 PT (ant)agonists of cell cycle regulation.
 XX
 PS Claim 1; Page 87-88; 115pp; English.
 XX
 CC R90533-R90556 are cyclin dependent kinase 4 (CDK4) binding proteins.
 CC encoded by clones of the plasmid pJG4-5-CDKBP. CDK4 binding proteins
 CC (CDK4-BP) may be used in an assay for screening test compounds as
 CC inhibitors of CDK/CDK4-BP interaction. The complexes formed by CDK4
 CC and p-type cyclins are strongly implicated in the control of the early
 CC G1 phase of the cell cycle and are strong candidates for controlling
 CC and/or preventing tumourigenesis and the onset of cancer. Nucleic acids
 CC encoding CDK4-BP or fragments of these may be used as probes/primers
 CC to diagnose the presence or absence of genetic lesions in a gene
 CC encoding 1 of the 24 CDK4-BP, and hence to diagnose the risk for a
 CC subject of developing a cell-proliferation associated disorder (e.g.
 CC cancer).
 XX
 SQ Sequence 382 AA;

Query Match 56.1%; Score 37; DB 17; Length 382;
 Best Local Similarity 70.0%; Pred. No. 75;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 INVSIRKKLR 14
 ||| : ||| : |
 Db 50 invavkklr 59

RESULT 8
 ID Y42419 standard; protein; 408 AA.

XX Y42419;
 AC
 DT 02-DEC-1999 (first entry)
 DE Jun N-terminal kinases (JNK)2 mutant.
 DE mitogen activated protein; MAP kinase; apoptosis; cancer; inflammation;
 KW intracellular signal transduction pathway; inhibitor; wildtype; ERK;
 KW extracellular signal regulated kinase; pyridinyl-imidazole.
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 108
 FT /note= "Met can be substituted by an amino acid
 FT with a small side chain, for example
 FT alanine or threonine"
 XX
 XX WO9942592-A1.
 XX 26-AUG-1999.
 XX 16-FEB-1999; 99WO-US03181.
 XX 18-FEB-1998; 98US-0025580.
 XX (VERT-) VERTEX PHARM INC.
 XX Su MS, Fox E, Wilson KP, Germann UA;
 XX WPI; 1999-540310/45.
 XX Method of designing Ser/Thr or Tyr kinase inhibitor useful for
 PT treating, example breast cancer, restenosis, asthma or hypertension -
 PT Example 6; Page 58-59; 71pp; English.
 XX This is the amino acid sequence of JNK2. Substitution of Methionine
 CC with an amino acid with a small side chain, such as Ala and Thr
 CC will create a mutant kinase which can theoretically bind to a
 CC pyridinyl-imidazole inhibitor of p38.
 CC (MAP) 1 kinase. MAP 1 kinases mediate intracellular signal transduction
 CC pathways and so have a role in many diverse human diseases. For example,
 CC kinases have been implicated in cell entry into apoptosis, cancer,
 CC Alzheimer's disease, angiotensin II and hematopoietic cytokine receptor
 CC signal transduction, oncoprotein signalling and mitosis, inflammation
 CC and infection, etc.
 CC Members of the MAP kinase family share sequence similarity and
 CC conserved structural domains, and include the extracellular-signal
 CC regulated kinases (ERKs), Jun N-terminal kinases (JNKs) and p38 kinases.
 CC The invention relates to methods for designing inhibitors of
 CC serine/threonine kinases, particularly MAP kinases, and tyrosine kinases
 CC through the use of ATP-binding site mutants of these kinases. The
 CC methods of this invention take advantage of the fact that the mutant
 CC kinases are capable of binding inhibitory compounds of other kinases
 CC with greater affinity than the corresponding wild-type kinase.
 XX Sequence 408 AA;
 XX
 XX Query Match 56.1%; Score 37; DB 20; Length 408;
 XX Best Local Similarity 70.0%; Pred. No. 80;
 XX Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 5 INVSIIKLKR 14
 DB 50 invavkklr 59
 RESULT 9
 R89408
 ID R89408 standard; Protein; 423 AA.

XX R89408;
 AC
 DT 26-APR-1996 (first entry)
 DE Stress activated protein kinase p54 alpha-I.
 DE Stress activated protein kinase; SAPK; p54 alpha-I; antibody;
 KW inflammation; hypoxia; heat stress.
 XX Rattus sp.
 OS
 XX CA2148898-A.
 XX 10-NOV-1995.
 XX 08-MAY-1995; 95CA-2148898.
 XX 09-MAY-1994; 94US-0240014.
 XX (GEHO) GEN HOSPITAL CORP.
 XX (ONTA-) ONTARIO CANCER INST.
 XX Avtuch J, Banerjee P, Kyriakis JM, Woodgett JR;
 XX WPI; 1996-106355/12.
 XX N-PSDB; T10638.
 XX DNA encoding recombinant p54 stress activated protein kinase - and
 PT related antibodies, useful for treating inflammation, hypoxia and
 PT heat stress, and for drug screening
 XX Claim 2; Page 34-35; 68pp; English.
 XX A new family of proline-directed stress-activated protein kinases
 CC (SAPK), designated p54 alpha-I, alpha-II, beta-I, beta-II and
 CC gamma (R89408-12, respectively), are products of a set of cDNA
 CC clones (see T10638-42) isolated from the brain cDNA library of
 CC a cycloheximide-treated rat. The SAPK p54 proteins are strongly
 CC activated by heat shock, tumour necrosis factor, interleukin-1-beta,
 CC sphingomyelinase, chemical protein synthesis inhibitors and
 CC ischaemia. They have potential utility in the modulation of the
 CC inflammatory response and the up-regulation of repair or
 CC protective cellular proteins following injury or chemical insult.
 XX Sequence 423 AA;
 XX
 XX Query Match 56.1%; Score 37; DB 17; Length 423;
 XX Best Local Similarity 70.0%; Pred. No. 83;
 XX Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 5 INVSIIKLKR 14
 DB 50 invavkklr 59
 RESULT 10
 R89409
 ID R89409 standard; Protein; 423 AA.
 XX
 XX R89409;
 XX 26-APR-1996 (first entry)
 XX Stress activated protein kinase p54 alpha-II.
 DE Stress activated protein kinase; SAPK; p54 alpha-II; antibody;
 KW inflammation; hypoxia; heat stress.
 XX Rattus sp.
 OS
 XX CA2148898-A.
 XX

```

XX      10-NOV-1995.
XX PD
XX XX
XX PF    08-MAY-1995;   95CA-2148898.
XX XX
XX PR    09-MAY-1994;   94US-0240014.
XX XX
XX PA    (GEHO ) GEN HOSPITAL CORP.
XX PA    (ONTA-) ONTARIO CANCER INST.
XX PI
XX AV     Avruch J, Banerjee P, Kyriakis JM, Woodgett JR;
XX WPI; 1996-106355/12.
XX N-PSDB; T10639.
XX XX
XX PT DNA encoding recombinant p54 stress activated protein kinase - and
XX PT related antibodies, useful for treating inflammation, hypoxia and
XX PT heat stress, and for drug screening
XX XX
XX CL     Claim 4; Page 35-36; 68pp; English.
XX XX
XX CC A new family of proline-directed stress-activated protein kinases
XX CC (SAPK), designated p54 alpha-I, alpha-II, beta-I, beta-II and
XX CC gamma (R99408-12, respectively), are products of a set of cDNA
XX CC clones (see T10638-42) isolated from the brain cDNA library of
XX CC a cycloheximide-treated rat. The SAPK p54 proteins are strongly
XX CC activated by heat shock, tumour necrosis factor, interleukin-1-beta,
XX CC sphingomyelinase, chemical protein synthesis inhibitors and
XX CC ischaemia. They have potential utility in the modulation of the
XX CC inflammatory response and the up-regulation of repair or
XX CC protective cellular proteins following injury or chemical insult.
XX XX
XX SQ     Sequence      423 AA;

Query Match          56.1%; Score 37; DB I7; Length 423;
Best Local Similarity 70.0%; Pred. No. 83;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      5 INVSIRKLKR 14
       |||::||| |
DB      50 invavkklsr 59

RESULT 11
Y34467
ID      Y34467 standard; Protein; 443 AA.
XX AC
XX AC      Y34467;
XX XX
XX DT      25-AUG-1999 (first entry)
XX XX
XX DE      Porphyromonas gingivalis protein PG122.
XX XX
XX KW      Porphyromonas gingivalis; PG: periodontal disease; gingivitis;
KW vaccine; antigenic.
XX OS      Porphyromonas gingivalis.
XX PN      WO9929870-A1.
XX PD      17-JUN-1999.
XX PF      10-DEC-1998;   98WO-AU01023.
XX PR      04-AUG-1998;   98AU-0005028.
PR      10-DEC-1997;   97AU-0000839.
PR      31-DEC-1997;   97AU-0001182.
PR      30-JAN-1998;   98AU-0001546.
PR      10-MAR-1998;   98AU-0002264.
PR      09-APR-1998;   98AU-0002911.
PR      23-APR-1998;   98AU-0003128.
PR      05-MAY-1998;   98AU-0003338.
```

PI Ross BC, Rothel LJ, Webb EA;
 XX WPI; 1999-385613/32.
 DR N-PSDB; X91561.
 XX Antigenic Porphyromonas gingivalis peptides for preventing
 PT gingivitis
 XX Claim 1; Page 303; 588pp; English.
 XX
 CC X91536 to X91801 encode two hundred and sixty six antigenic
 CC Porphyromonas gingivalis (PG) polypeptide sequences given in Y34318 to
 CC Y34583. X91802 to X91989 represent PCR primers used in the isolation
 CC of the PG polypeptides. The PG polypeptides have antibacterial activity
 CC with a vaccine mechanism of action. The PG polypeptides can be used as
 CC vaccines especially against Porphyromonas gingivalis. Probes can be
 CC used to detect Porphyromonas gingivalis in standard hybridisation assays.
 CC Porphyromonas gingivalis is involved in periodontal disease especially
 CC gingivitis.
 XX
 SQ Sequence 451 AA;
 Query Match 56.1%; Score 37; DB 20; Length 451;
 Best Local Similarity 61.5%; Pred. No. 89;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 2 ERIINVSIIKKLR 14
 III: I: IIII
 DB 294 erivtvgkklsr 306
 RESULT 13
 R76062
 ID R76062 standard; Protein; 540 AA.
 AC R76062;
 XX
 DT 12-MAR-1996 (first entry)
 XX
 DE Protein kinase PKU beta.
 XX
 KW Protein kinase; PKU beta; diagnosis; treatment; coronary; disease;
 KW neuropathy; cancer; immune; endocrine; cell; division; growth;
 KW differentiation; apoptosis; development; canceration.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 15 /note= "corresponding codon AAT"
 FT Misc-difference 467 /note= "corresponding codon GCA"
 FT Misc-difference 482 /note= "corresponding codon GCA"
 XX
 PN JP07132093-A.
 XX
 PD 23-MAY-1995.
 XX
 PF 12-NOV-1993; 93JP-0306095.
 XX
 PR 12-NOV-1993; 93JP-0306095.
 XX
 XX (IDAT/) IDATE T.
 PA (MTK) MITSUI TOATSU CHEM INC.
 PA
 XX WPI; 1995-220116/29.
 DR N-PSDB; Q94050.
 XX
 XX New protein kinase(s), PKU alpha and PKU beta - and antibodies
 PT against them, for the treatment or prevention of coronary diseases,
 PT neuropathy, cancer, immune or endocrine diseases

XX Claim 5; Fig 1; 7pp; Japanese.
 XX
 CC Q94050 encodes R76062 protein kinase PKU beta. PKU beta is useful
 CC in the elucidation of cell division, development, growth, apoptosis,
 CC differentiation and canceration. The kinase, antibodies or the gene
 CC (including fragments, and sense and antisense RNA) can be used for
 CC the diagnosis and treatment of neuropathy, cancer and immune,
 CC endocrine and coronary diseases.
 XX
 SQ Sequence 540 AA;
 Query Match 56.1%; Score 37; DB 16; Length 540;
 Best Local Similarity 53.8%; Pred. No. 1.1e+02;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 2 ERIINVSIIKKLR 14
 III: I: IIII
 DB 199 ervnlhrelkr 211
 RESULT 14
 R60178
 ID R60178 standard; Protein; 776 AA.
 XX
 AC R60178;
 XX
 DT 03-APR-1995 (first entry)
 XX
 DE Lethal factor of Bacillus anthracis.
 XX
 KW Anthrax; Bacillus anthracis; fusion protein; lethal factor;
 KW protective antigen; cell killing; targeting; pathogen;
 KW intracellular; HIV; human immunodeficiency virus; toxin.
 XX
 OS Bacillus anthracis.
 XX
 PN W09418332-A.
 XX
 PD 18-AUG-1994.
 XX
 PF 14-FEB-1994; 94WO-US01624.
 XX
 PR 12-FEB-1993; 93US-0021601.
 PR 25-JUN-1993; 93US-0082849.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Arora N, Klimpel K, Leppia SH, Nichols PJ, Singh Y;
 DR WPI; 1994-279753/34.
 DR N-PSDB; Q70179.
 XX
 PT Nucleic acid encoding anthrax toxin fusion protein - useful for
 PT targeting toxin to specific cells, eg for killing tumour cells
 PT or HIV-infected cells
 XX
 PS Disclosure; Page 75-77; 124pp; English.
 XX
 CC The sequence encoding the lethal factor of Bacillus anthracis may be
 CC used in the construction of a nucleic acid which encodes a fusion
 CC protein comprising the anthrax protective antigen binding domain of
 CC the native anthrax lethal factor and a sequence encoding an activity
 CC inducing domain of a second protein. The fusion proteins are useful
 CC for the specific killing of tumour cells or the killing of cells
 CC infected with intracellular pathogens, especially HIV.
 XX
 SQ Sequence 776 AA;
 Query Match 56.1%; Score 37; DB 15; Length 776;
 Best Local Similarity 53.8%; Pred. No. 1.6e+02;

Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ERIINVSIIKKLR 14
Db 446 ervnlhrelkr 458

Search completed: May 23, 2001, 11:11:38
Job time: 181 sec

QY 1 NERLIINVSIIKKL 13
II: I: I: I: I: I: I:
Db 248 negeinlsleelk 260

RESULT 15

R76061
ID R76061 standard; Protein; 787 AA.

XX AC R76061;

XX DT 12-MAR-1996 (first entry)

XX DE Protein kinase PKU alpha.

XX KW Protein kinase; PKU alpha; diagnosis; treatment; coronary; disease;

XX KW neuropathy; cancer; immune; endocrine; cell; division; growth;

XX KW differentiation; apoptosis; development; canceration.

XX OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 250 /note= "corresponding codon ATC"

FT Misc-difference 264 /note= "corresponding codon CTC"

FT Misc-difference 290 /note= "corresponding codon AAG"

FT Misc-difference 483 /note= "corresponding codon GCG"

FT Misc-difference 484 /note= "corresponding codon TGT"

FT Misc-difference 713 /note= "corresponding codon GCC"

FT Misc-difference 728 /note= "corresponding codon GCC"

XX JF07132093-A.

XX PD 23-MAY-1995.

XX PF 12-NOV-1993; 93JP-0306095.

XX PR 12-NOV-1993; 93JP-0306095.

XX (IDAT/) IDATE T.

XX (MITK) MITSUI TOATSU CHEM INC.

XX WPI; 1995-220116/29.

XX N-PSDB; Q94049.

XX New protein kinase(s), PKU alpha and PKU beta - and antibodies

XX against them, for the treatment or prevention of coronary diseases,

XX neuropathy, cancer, immune or endocrine diseases

XX Claim 5; Fig 1; 7pp; Japanese.

XX Q94049 encodes R76061 protein kinase PKU alpha. PKU alpha is useful

XX in the elucidation of cell division, development, growth, apoptosis,

XX differentiation and canceration. The kinase, antibodies or the gene

XX (including fragments, and sense and antisense RNA) can be used for

XX the diagnosis and treatment of neuropathy, cancer and immune,

XX endocrine and coronary diseases.

XX SQ Sequence 787 AA;

Query Match 56.1%; Score 37; DB 16; Length 787;

Best Local Similarity 53.8%; Pred. No. 1.6e+02;

Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:09:43 ; Search time 58.85 Seconds
(without alignments)
4.570 Million cell updates/sec

Title: US-09-522-217-2_COPY_92_105

Perfect score: 66

Sequence: 1 NERIINVSIIKKLR 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PTCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	56.1	424	4	US-09-025-580-32
2	37	56.1	776	1	US-08-021-601-2
3	37	56.1	776	1	US-08-082-849B-2
4	37	56.1	776	5	PCT-US94-01624-2
5	35	53.0	218	4	US-09-081-689-6
6	35	53.0	233	2	US-08-771-783-2
7	35	53.0	260	4	US-08-081-929-10
8	35	53.0	270	1	US-08-452-052-1
9	35	53.0	373	1	US-08-015-986A-14
10	35	53.0	373	2	US-08-446-363-14
11	35	53.0	411	1	US-07-684-135A-2
12	35	53.0	238	1	US-08-015-973-1
13	35	53.0	238	2	US-08-448-164-1
14	35	53.0	238	4	US-08-081-929-2
15	34	51.5	102	2	US-08-767-096-4
16	34	51.5	239	3	US-08-286-819A-44
17	34	51.5	239	3	US-08-980-357-44
18	34	51.5	384	4	US-09-209-668-15
19	34	51.5	384	4	US-09-025-580-33
20	34	51.5	384	4	US-09-025-580-35
21	34	51.5	384	5	PCT-US94-08119-12
22	34	51.5	384	5	PCT-US94-12913A-12
23	34	51.5	422	4	US-09-025-580-3
24	34	51.5	422	4	US-09-025-580-34
25	34	51.5	427	4	US-09-025-580-36
26	34	51.5	464	4	US-09-025-580-28
27	34	51.5	914	1	US-08-484-105-2

28	34	51.5	914	1	US-08-484-106-2	Sequence 2, Appli
29	33	50.0	224	1	US-08-170-294-2	Sequence 2, Appli
30	33	50.0	224	2	US-08-664-855-2	Sequence 2, Appli
31	33	50.0	224	3	US-09-049-289-2	Sequence 2, Appli
32	33	50.0	393	2	US-08-837-593-2	Sequence 2, Appli
33	33	50.0	984	1	US-08-242-932-2	Sequence 2, Appli
34	33	50.0	984	1	US-08-714-481-2	Sequence 2, Appli
35	33	50.0	984	5	PCT-US95-06111-2	Sequence 2, Appli
36	32.5	49.2	383	4	US-08-844-188-38	Sequence 38, Appl
37	32	48.5	107	1	US-08-369-796-16	Sequence 16, Appl
38	32	48.5	107	2	US-08-852-091-16	Sequence 16, Appl
39	32	48.5	107	5	PCT-US95-17025-16	Sequence 16, Appl
40	32	48.5	165	4	US-09-087-465-30	Sequence 30, Appl
41	32	48.5	210	1	US-08-286-872-2	Sequence 2, Appli
42	32	48.5	382	4	US-09-025-580-29	Sequence 29, Appl
43	32	48.5	382	4	US-09-025-580-30	Sequence 30, Appl
44	32	48.5	383	1	US-08-597-545-2	Sequence 2, Appli
45	32	48.5	383	1	US-08-457-135-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-025-580-32
; Sequence 32, Application US/09025580
; Patent No. 6162613
; GENERAL INFORMATION:
; APPLICANT: Su, Michael Shin-San
; APPLICANT: Fox, Ted
; APPLICANT: Wilson, Keith Phillip
; APPLICANT: Germann, Ursula A.
; TITLE OF INVENTION: Methods For Designing Inhibitors of
; TITLE OF INVENTION: Serine/Threonine Kinases and Tyrosine Kinase
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,580
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley, James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: VPI 97-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-025-580-32

Query Match 56.1%; Score 37; DB 4; Length 424;
Best Local Similarity 70.0%; Pred. No. 57;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 5 INVSIIKKLR 14

```

Db      50 INVAVKLSR 59
      III::III I
      STREET: Two Embarcadero Center, Eighth Floor
      CITY: San Francisco
      STATE: California
      COUNTRY: USA
      ZIP: 94111-3834
      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.30
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/082,849B
      FILING DATE: 25-JUN-1993
      CLASSIFICATION: 514
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 08/021,601
      FILING DATE: 12-FEB-1993
      ATTORNEY/AGENT INFORMATION:
      NAME: Weber, Kenneth A.
      REGISTRATION NUMBER: 31,677
      REFERENCE/DOCKET NUMBER: 15280-161-1
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (415) 576-0200
      TELEFAX: (415) 576-0300
      INFORMATION FOR SEQ ID NO: 2:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 776 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      US-08-082-849B-2

      Query Match      56.1%; Score 37; DB 1; Length 776;
      Best Local Similarity 53.8%; Pred. No. 1e+02;
      Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      1 NERIINVSIIKLIK 13
      II::II::II
Db      248 NEQEINLSLEELK 260
      II::II::II

RESULT 4
PCT-US94-01624-2
; Sequence 2, Application PC/TUS9401624
; GENERAL INFORMATION:
; APPLICANT: Leppla, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW
; STREET: Steuart Street Tower, 20th Floor, One Market
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01624
; FILING DATE: June 25, 1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.

Db      50 INVAVKLSR 59
      III::III I
      STREET: 133 Carnegie Way, Suite 400
      CITY: Atlanta
      STATE: Georgia
      COUNTRY: USA
      ZIP: 30303
      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.25
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/021,601
      FILING DATE: 19930212
      CLASSIFICATION: 514
      ATTORNEY/AGENT INFORMATION:
      NAME: Spratt, Gwendolyn D.
      REGISTRATION NUMBER: 36,016
      REFERENCE/DOCKET NUMBER: 1414.057
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: 404/688-0770
      TELEFAX: 404/688-9880
      INFORMATION FOR SEQ ID NO: 2:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 776 amino acids
      TYPE: AMINO ACID
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      US-08-021-601-2

      Query Match      56.1%; Score 37; DB 1; Length 776;
      Best Local Similarity 53.8%; Pred. No. 1e+02;
      Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY      1 NERIINVSIIKLIK 13
      II::II::II
Db      248 NEQEINLSLEELK 260
      II::II::II

RESULT 3
US-08-082-849B-2
; Sequence 2, Application US/08082849B
; Patent No. 5677274
; GENERAL INFORMATION:
; APPLICANT: Leppla, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
; TITLE OF INVENTION: Related Methods
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP

```

REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-115
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 776 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-01624-2

Query Match 56.1%; Score 37; DB 5; Length 776;
Best Local Similarity 53.8%; Pred. No. 1e+02;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 NERIINVSIIKKL 13
DB 248 NQEQINLSUEEL 260

RESULT 5
US-09-081-689-6
Sequence 6, Application US/09081689
Patent No. 6165992
GENERAL INFORMATION:
APPLICANT: Wallis, Nicola G.
APPLICANT: Zalacain, Magdalena
APPLICANT: Throup, John
APPLICANT: Biswas, Sanjoy
TITLE OF INVENTION: Histidine Kinase
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/081,689
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/048,347
FILING DATE: 30-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Falk, Stephen T
REGISTRATION NUMBER: 36,795
REFERENCE/DOCKET NUMBER: GM10009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2488
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 218 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

Query Match 53.0%; Score 35; DB 4; Length 218;
Best Local Similarity 46.2%; Pred. No. 63;
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 2 ERIINVSIIKKLR 14
DB 186 DRVIDVHIKLR 198

RESULT 6
US-08-771-783-2
Sequence 2, Application US/08771783
Patent No. 5962259
GENERAL INFORMATION:
APPLICANT: Wallis, Nicola G.
APPLICANT: Hodgson, John E.
TITLE OF INVENTION: No. 5962259el Compounds
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-2799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/771,783
FILING DATE: 20-DEC-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9526332.3
FILING DATE: 22-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9526358.8
FILING DATE: 22-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9526361.2
FILING DATE: 22-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31313
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 233 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-771-783-2

Query Match 53.0%; Score 35; DB 2; Length 233;
Best Local Similarity 42.9%; Pred. No. 68;
Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 NERIINVSIIKKLR 14
DB 195 DERIVDVHIKLRQ 208

RESULT 7
US-08-081-929-10
Sequence 10, Application US/08081929
Patent No. 6160090
GENERAL INFORMATION:
APPLICANT: Schlessinger, Joseph
APPLICANT: Barnea, Gilad

APPLICANT: Grumet, Martin H.
APPLICANT: Margolis, Richard U.
TITLE OF INVENTION: A NEW CLASS OF RPTPASS: THEIR
STRUCTURAL DOMAINS AND LIGANDS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/081,929
FILING DATE: 23-JUN-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30742
REFERENCE/DOCKET NUMBER: 7683-041-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 260 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-081-929-10

Query Match 53.0%; Score 35; DB 4; Length 260;
Best Local Similarity 66.7%; Pred. No. 76;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 INVSIIKLLK 13
DB 43 VNVNKKLK 51

RESULT 8
US-08-452-052-1
Sequence 1, Application US/08452052
Patent No. 5766922
GENERAL INFORMATION:
APPLICANT: Peles, Elior
TITLE OF INVENTION: FUNCTIONAL LIGANDS FOR THE AXONAL CELL
RECOGNITION MOLECULE CONTACTIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/452,052
FILING DATE: 26-MAY-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-075
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 270 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-452-052-1

Query Match 53.0%; Score 35; DB 1; Length 270;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 INVSIIKLLK 13
DB 43 VNVNKKLK 51

RESULT 9
US-08-015-986A-14
Sequence 14, Application US/08015986A
Patent No. 5532123
GENERAL INFORMATION:
APPLICANT: Schlessinger, Joseph
TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
TITLE OF INVENTION: PHOSPHATASE-GAMMA
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/015,986A
FILING DATE: 10-FEB-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,972
REFERENCE/DOCKET NUMBER: 7683-028
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-015-986A-14

Query Match 53.0%; Score 35; DB 1; Length 373;
Best Local Similarity 66.7%; Pred. No. 11e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

QY 5 INVSIIKKLK 13
DB 40 VNVNKKLK 48

RESULT 10
US-08-446-363-14
; Sequence 14, Application US/08446363
; Patent No. 5891700
; GENERAL INFORMATION:
; APPLICANT: Schlusser, Joseph
; TITLE OF INVENTION: NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-GAMMA
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,363
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/015,986
; FILING DATE: 10-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-446-363-14

```

```

Query Match 53.0%; Score 35; DB 2; Length 373;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 5 INVSIIKKLK 13
DB 40 VNVNKKLK 48

RESULT 11
US-07-684-135A-2
; Sequence 2, Application US/07684135A
; Patent No. 5243039
; GENERAL INFORMATION:
; APPLICANT: Schendel, Frederick J. and
; APPLICANT: Flickinger, M. C.
; TITLE OF INVENTION: Bacillus MGA 3 Aspartokinase
; TITLE OF INVENTION: II Gene
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould

```

```

; STREET: 3100 No. 5243039west Center
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb
; COMPUTER: No. 5243039thgate 386
; OPERATING SYSTEM: DOS 4.0
; SOFTWARE: WordPerfect- 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/684,135A
; FILING DATE: 19910412
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D.
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 600.216-US-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (612) 332-5300
; TELEFAX: (612) 332-9081
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: Linear
; MOLECULE TYPE: Polypeptide
; FEATURE:
; NAME/KEY: Aspartokinase II -- dimer subunit
; LOCATION: 1 to 411
US-07-684-135A-2

```

```

Query Match 53.0%; Score 35; DB 1; Length 411;
Best Local Similarity 35.7%; Pred. No. 1.2e+02;
Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 NERIINVSIIKKLK 14
DB 110 NARILNIEFKIQK 123

RESULT 12
US-08-015-973-1
; Sequence 1, Application US/08015973
; Patent No. 5604094
; GENERAL INFORMATION:
; APPLICANT: Schlusser, Joseph
; TITLE OF INVENTION: A NOVEL RECEPTOR-TYPE PHOSPHOTYROSINE
; TITLE OF INVENTION: PHOSPHATASE-BETA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/015,973
; FILING DATE: 10-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090

```

QY 5 INVSIIKKLK 13
:|:|:|:|:|
Db 73 VNVNKKLK 81

STATE: PA
COUNTRY: USA

```

;
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/767,096
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Schreck, Patricia A.
; REGISTRATION NUMBER: 33,777
; REFERENCE/DOCKET NUMBER: ATG50033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5031
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 102 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-767-096-4

```

```

Query Match          51.5%; Score 34; DB 2; Length 102;
Best Local Similarity 60.0%; Pred. No. 44;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 3 RIINVSIKKL 12
   |::|||::||
Db 20 RVLNVSEKL 29

```

Search completed: May 23, 2001, 11:09:44
Job time: 67 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:12:53 ; Search time 70.54 Seconds
(without alignments)
13.639 Million cell updates/sec

Title: US-09-522-217-2_COPY_92_105

Perfect score: 66

Sequence: 1 NERIINVSIIKLKR 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_67:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	65.2	410	1 ZEBP4L	Ea47 protein - pha
2	41	62.1	615	2 T06108	hypothetical prote
3	41	62.1	908	2 F64207	phosphotransferase
4	40	60.6	276	2 C82420	lipase-related pro
5	40	60.6	292	2 T32000	hypothetical prote
6	40	60.6	322	2 G64151	hypothetical prote
7	40	60.6	489	2 T47747	cysteine-trna liga
8	40	60.6	1228	2 S60085	nitrate reductase
9	39	59.1	210	2 D81311	thiamin-phosphate
10	39	59.1	223	2 E69747	two-component resp
11	39	59.1	417	2 C70391	hypothetical prote
12	39	59.1	572	2 T45139	chaperone protein
13	39	59.1	788	2 S67595	hypothetical prote
14	38	57.6	208	2 C69552	DNA repair protein
15	38	57.6	333	2 T11212	3-oxoacyl-lacyl-ca
16	38	57.6	702	2 S59428	probable membrane
17	38	57.6	1245	2 D71613	GAF domain protein
18	38	57.6	4092	1 S38128	dynamin heavy chain
19	37	56.1	171	2 F82778	hypothetical prote
20	37	56.1	283	2 H69179	conserved hypothet
21	37	56.1	334	2 D27335	hypothetical nox3
22	37	56.1	382	2 JC3531	c-Jun amino-termin
23	37	56.1	391	2 T20752	hypothetical prote
24	37	56.1	423	2 S43967	p54-alpha stress-a
25	37	56.1	423	2 S43968	p54-alpha-2 stress
26	37	56.1	424	2 A53480	c-Jun amino-termin
27	37	56.1	620	1 F64410	molybdenum cofacto
28	37	56.1	741	2 A64524	endopeptidase Clp
29	37	56.1	742	2 C71983	endopeptidase Clp

ALIGNMENTS

RESULT 1

ZEBP4L

Ea47 protein - phage lambda

C:Species: phage lambda

C>Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 23-Jul-1999

C:Accession: E43009; B43015; A04377

R:Daniels, D.

submitted to the Nucleic Acid Sequence Database, September 1982

A:Reference number: A94614

A:Accession: E43009

A:Molecule type: DNA

A:Residues: 1-410 <DAN>

R:Sanger, F.; Coulson, A.R.; Hong, G.F.; Hill, D.F.; Petersen, G.B.

J. Mol. Biol. 162, 729-773, 1982

A:Title: Nucleotide sequence of bacteriophage lambda DNA.

A:Reference number: A92891; MUID:83189071

A:Accession: B43015

A:Molecule type: DNA

A:Residues: 1-410 <SAN>

A:Cross-references: GB:J02459; GB:M17233; GB:M24325; GB:V00636; GB:X00906; NID:921510

C:Genetics:

A:Gene: Ea47

A:Map position: 49.31-46.78

C:Superfamily: phage lambda Ea47 protein

Query Match 65.2%; Score 43; DB 1; Length 410;
Best Local Similarity 57.1%; Pred. No. 5.7;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NERIINVSIIKLKR 14

||::||::||::||

Db 117 NESMLNIGIKKLVR 130

RESULT 2

T06108

hypothetical protein T5J17.190 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 22-Oct-1999

C:Accession: T06108

R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.

submitted to the Protein Sequence Database, March 1999

A:Reference number: Z15184

A:Accession: T06108

A:Molecule type: DNA

A:Residues: 1-615 <BEV>

A:Cross-references: EMBL:AL035708; GSPDB:GN00062; ATSP:T5J17.190

A:Experimental source: cultivar Columbia; BAC clone T5J17

C:Genetics:

A:Gene: ATSP:T5J17.190

A:Map position: 4

Query Match 62.1%; Score 41; DB 2; Length 615;
 Best Local Similarity 57.1%; Pred. No. 19;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NERIINVSIIKKLR 14
 || : | : |||||
 Db 350 NEAVANDNIKKLKK 363

RESULT 3
 F64207
 phosphotransferase enzyme II ABC component (ptsG) homolog - Mycoplasma genitalium
 C:Species: Mycoplasma genitalium
 C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 21-Jul-2000
 C:Accession: F64207
 R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
 M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.
 C.A.; Venter, J.C.
 Science 270, 397-403, 1995
 A:Title: The minimal gene complement of Mycoplasma genitalium.
 A:Reference number: A64200; MUID:96026346
 A:Accession: F64207
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-908 <TIGR>
 A:Cross-references: GB:U39687; GB:L43967; NID:g3844663; PIDN:AAC71287.1; PID:g1045745;
 A:Experimental source: strain G-37
 C:Genetics:
 A:Genetic code: SGC3
 C:Superfamily: phosphotransferase system glucose-specific enzyme II, factor III homolog
 F:739-895/Domain: phosphotransferase system glucose-specific enzyme II, factor III homolog

Query Match 62.1%; Score 41; DB 2; Length 908;
 Best Local Similarity 50.0%; Pred. No. 29;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NERIINVSIIKKLR 14
 : | : | : |||||
 Db 850 SESVVEVDLKKLR 863

RESULT 4
 C82420
 lipase-related protein VCA0754 [imported] - Vibrio cholerae (group O1 strain N16961)
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
 C:Accession: C82420
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, B.
 L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833
 A:Accession: C82420
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-276 <HEI>
 A:Cross-references: GB:AE004404; GB:AE003853; NID:g9659174; PIDN:AAF96652.1; GSPDB:GN001
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VCA0754
 A:Map position: 2

Query Match 60.6%; Score 40; DB 2; Length 276;
 Best Local Similarity 66.7%; Pred. No. 13;
 Matches 8; Conservative 3; Mismatches 3; Indels 1; Gaps 0;

Qy 3 RIINVSIIKKLR 14
 ||| : | : | |||

Db 9 RIIDVTVKPLKR 20

RESULT 5
 T32000
 hypothetical protein B0281.8 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 02-Sep-2000
 C:Accession: T32000
 R:Pauley, A.; Scheet, P.
 submitted to the EMBL Data Library, July 1997
 A:Description: The sequence of C. elegans cosmid B0281.
 A:Reference number: Z21109
 A:Accession: T32000
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-292 <PAU>
 A:Cross-references: EMBL:AF016666; PIDN:AAB66086.1; GSPDB:GN00020; CESP:B0281.8
 A:Experimental source: strain Bristol N2; clone B0281
 C:Genetics:
 A:Gene: CESP:B0281.8
 A:Map position: 2
 A:Introns: 27/1; 85/3; 108/3; 171/2; 186/3; 270/1
 C:Superfamily: RING finger homolog
 F:2-57/Domain: RING finger homolog <RRN>

Query Match 60.6%; Score 40; DB 2; Length 292;
 Best Local Similarity 38.5%; Pred. No. 14;
 Matches 5; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NERIINVSIIKKLR 13
 : | : | : | : |||
 Db 176 SEOLLNKKVKKIK 188

RESULT 6
 G64151
 hypothetical protein HI0412 - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 08-Oct-1999
 C:Accession: G64151
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
 Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
 D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630
 A:Accession: G64151
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-322 <TIGR>
 A:Cross-references: GB:U32724; GB:L42023; NID:g1573378; PIDN:AAC22071.1; PID:g1573385
 A:Note: best homolog was a hypothetical protein from Escherichia coli
 C:Superfamily: conserved hypothetical protein HI0176

Query Match 60.6%; Score 40; DB 2; Length 322;
 Best Local Similarity 66.7%; Pred. No. 15;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NERIINVSIIKKLR 12
 || : ||| | : | |
 Db 5 NEKIINSSVKML 16

RESULT 7
 T47747
 cytochrome-c oxidase - Arabidopsis thaliana
 N:Alternate names: protein F18021.260
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 19-May-2000

C:Accession: T47747
 R:Benes, V.; Wurmbach, E.; Drzonek, H.; Ansoorge, W.; Meves, H.W.; Rudd, S.; Lemcke, K.; submitted to the Protein Sequence Database, April 2000
 A:Reference number: 224474
 A:Accession: T47747
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-489 <BEN>
 A:Cross-references: EMBL:ATF18021
 A:Experimental source: cultivar Columbia; BAC clone F18021
 C:Genetics:
 A:Map position: 3
 A:Introns: 61/2; 82/3; 132/3; 189/3; 212/2; 271/3; 310/1; 340/1
 A:Note: F18021.260
 C:Superfamily: cysteine--tRNA ligase

Query Match 60.6%; Score 40; DB 2; Length 489;
 Best Local Similarity 66.7%; Pred. No. 23;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 RIINVSIRKLK 14
 : ||||| :
 Db 343 KFINVSIRKLK 354

RESULT 8
 S60085
 N:Alternate names: respiratory nitrate reductase alpha chain narG
 C:Species: Bacillus subtilis
 C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 20-Jun-2000
 C:Accession: S60085; I40426; I40430; S52915; E69664; S60574
 R:Crúz Ramos, H.; Boursier, L.; Moszer, I.; Kunst, F.; Danchin, A.; Glaser, P.
 EMBO J. 14, 5984-5994, 1995
 A:Title: Anaerobic transcription activation in Bacillus subtilis: identification of distal
 A:Reference number: S60080; MUID:96112813
 A:Accession: S60085
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1228 <CRU>
 A:Cross-references: EMBL:249884; NID:g971335; PIDN:CAA90045.1; PID:g971341
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995
 R:Hoffmann, T.; Troup, B.; Szabo, A.; Hungerer, C.; Jahn, D.
 FEBS Microbiol. Lett. 131, 219-225, 1995
 A:Title: The anaerobic life of Bacillus subtilis: cloning of the genes encoding the resp
 A:Reference number: S60573; MUID:96004574
 A:Accession: I40426
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-46,'d',48-1228 <RE2>
 A:Cross-references: EMBL:X91819; NID:g1009365; PIDN:CAA62926.1; PID:g1009366
 A:Accession: I40430
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 567-1228 <RES>
 A:Cross-references: EMBL:X85014; NID:g1001847; PIDN:CAA59371.1; PID:g1001848
 R:Hoffmann, T.; Hungerer, C.; Troup, B.; Jahn, D.
 submitted to the EMBL Data Library, February 1995
 A:Description: The anaerobic life of Bacillus subtilis: cloning and regulation of the na
 A:Reference number: S52915
 A:Accession: S52915
 A:Molecule type: DNA
 A:Residues: 567-739,'n',741-777,779-880,'C',882-1174,'RPY',1179-1190,'NRS',1194-1228 <HC
 A:Cross-references: EMBL:X85014
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
 C.; Bron, S.; Brouillet, S.; Bruchli, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Galler
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Königstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadie, Y.; Sato, T.; Scanl
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Segiguchi, J.; Sekowska, A.; Se
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
 A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili
 A:Reference number: A69580; MUID:98044033
 A:Accession: E69664
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: narG
 C:Superfamily: nitrate reductase alpha chain
 C:Keywords: 4Fe-4S; electron transfer; iron-sulfur protein; membrane-associated compl
 F;58,62,97/Binding site: iron-sulfur clusters (Cys) (covalent) #status predicted

Query Match 60.6%; Score 40; DB 2; Length 1228;
 Best Local Similarity 61.5%; Pred. No. 59;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NERIINVSIRKLK 13
 : : : : :
 Db 1209 NQDLNVIRKLK 1221

RESULT 9
 D81311
 A:Title: Thiophosphate pyrophosphorylase (EC 2.5.1.3) Cj1081c [imported] - Campylobacter j
 C:Species: Campylobacter jejuni
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 26-May-2000
 C:Accession: D81311
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chil
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Ba
 Nature 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
 A:Reference number: A81250; MUID:20150912
 A:Accession: D81311
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-210 <PAR>
 A:Cross-references: GB:AL139077; GB:AL111168; NID:g968444; PIDN:CA873336.1; PID:g9696
 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 A:Gene: thiE; Cj1081c
 C:Superfamily: probable thiamin-phosphate pyrophosphorylase; thiamin-phosphate pyroph
 C:Keywords: transferase

Query Match 59.1%; Score 39; DB 2; Length 210;
 Best Local Similarity 50.0%; Pred. No. 15;
 Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 NERIINVSIRKLK 14
 : : : : :
 Db 103 DEKIIGLSLAKLEQ 116

RESULT 10
 E69747
 C:Species: Bacillus subtilis
 A:Title: two-component response regulator [YbdK] homolog ybdJ - Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 C:Accession: E69747
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
 C.; Bron, S.; Brouillet, S.; Bruchli, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gal

Iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron, akeuchi, M.; Tamakoshi, A.; Tanaka, R.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A:Reference number: A69580; MUID:98044033
 A:Accession: F69747
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-223 <KUN>
 A:Cross-references: GB:Z99105; GB:AL009126; NID:g2632457; PIDN:CAB11994.1; PID:g2632486
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: YbdJ
 C:Superfamily: ompR protein; response regulator homology
 C:Keywords: phosphoprotein
 F:6-112/Domain: response regulator homology <RRH>
 F:52/Binding site: phosphate (Asp) (covalent) #status predicted

 Query Match 59.1%; Score 39; DB 2; Length 223;
 Best Local Similarity 53.8%; Pred. No. 16;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

 Qy 1 NERIINVSIIKKL 13
 | : | | | | : |
 Db 184 NPNIVNVHIKKIR 196

 RESULT 11
 C70391
 hypothetical protein aq_1059 - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
 C:Accession: C70391
 R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.
 A:Reference number: A70300; MUID:98196666
 A:Accession: C70391
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-417 <AQF>
 A:Cross-references: GB:AE000721; NID:g2983544; PIDN:AAC07128.1; PID:g2983554; GB:AE00065
 A:Experimental source: strain VF5
 C:Genetics:
 A:Gene: aq_1059

 Query Match 59.1%; Score 39; DB 2; Length 417;
 Best Local Similarity 50.0%; Pred. No. 30;
 Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

 Qy 2 ERIINVSIIKKL 13
 | : | : | : | : |
 Db 102 EKLIDISVKELK 113

 RESULT 12
 T45139
 chaperone protein thsB [imported] - *Pyrodicticum occultum*
 C:Species: *Pyrodicticum occultum*
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000
 C:Accession: T45139
 R:Frey, G.
 submitted to the EMBL Data Library, June 1998
 A:Reference number: 22924

A:Accession: T45139
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-572 <FRE>
 A:Cross-references: EMBL:AJ006550; PIDN:CAA07096.1
 C:Genetics:
 A:Gene: thsB
 C:Superfamily: molecular chaperone t-complex-type

 Query Match 59.1%; Score 39; DB 2; Length 572;
 Best Local Similarity 46.2%; Pred. No. 41;
 Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

 Qy 2 ERIINVSIIKKLR 14
 | : | | | : | : |
 Db 295 EKINVAVERMKR 307

 RESULT 13
 S67595
 hypothetical protein YDL060w - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: hypothetical protein D2544
 C:Species: *Saccharomyces cerevisiae*
 C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 20-Jun-2000
 C:Accession: S67595
 R:Blöcker, H.; Brandt, P.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S67587
 A:Accession: S67595
 A:Molecule type: DNA
 A:Residues: 1-788 <BL0>
 A:Cross-references: EMBL:Z74108; NID:g1431062; PID:g1431063; GSPDB:GN00004; MIPS:YDL0
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: MIPS:YDL060w
 A:Map position: 4L

 Query Match 59.1%; Score 39; DB 2; Length 788;
 Best Local Similarity 61.5%; Pred. No. 56;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

 Qy 1 NERIINVSIIKKL 13
 | : | | | | | : |
 Db 120 SKRIENVHIKKFK 132

 RESULT 14
 C69552
 DNA repair protein homolog - *Archaeoglobus fulgidus*
 C:Species: *Archaeoglobus fulgidus*
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: C69552
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod
 :; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E
 Glöck, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes,
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
 A:Reference number: A69250; MUID:98049343
 A:Accession: C69552
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-208 <KLE>
 A:Cross-references: GB:AE001108; GB:AE000782; NID:g2689431; PIDN:AAB91244.1; PID:g265

 Query Match 57.6%; Score 38; DB 2; Length 208;
 Best Local Similarity 61.5%; Pred. No. 22;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ERIINVSIIKKLR 14
:||| | || :|||
Db 179 QRIANASIAELKR 191

RESULT 15

Tl1212
3-oxoacyl-[acyl-carrier-protein] synthase homolog - Streptomyces glaucescens
C:Species: Streptomyces glaucescens
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-Aug-1999
C:Accession: Tl1212
R:Summers, R.G.; Ali, A.; Shen, B.; Wessel, W.A.; Hutchinson, C.R.
Biochemistry 34, 9389-9402, 1995
A:Title: Malonyl-coenzyme A:acyl carrier protein acyltransferase of Streptomyces glaucescens
A:Reference number: Z17254; MUID:95352622
A:Accession: Tl1212
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-333 <SUM>
A:Cross-references: EMBL:L43074; NID:g870805; PID:g870807
C:Genetics:
A:Gene: fabH
C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase III

Query Match 57.6%; Score 38; DB 2; Length 333;
Best Local Similarity 61.5%; Pred. No. 35;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NERIINVSIIKKLR 13
||||| :| |
Db 261 NERIIDSMVTKL 273

Search completed: May 23, 2001, 11:12:54
Job time: 257 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:20:26 ; Search time: 40.06 Seconds
(without alignments)
11.971 Million cell updates/sec

Title: US-09-522-217-2_COPY_92_105
Perfect score: 66
Sequence: 1 NERIINVSIIKLKR 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	65.2	410	1 VE47_LAMB	P03752 bacterioph
2	41	62.1	908	1 PTGA_MYCGE	P47315 mycoplasma
3	40	60.6	322	1 RLUC_HAEIN	P44433 haemophilus
4	40	60.6	1228	1 NARG_BACSU	P42175 bacillus su
5	38	57.6	911	1 H101_ARATH	P42170 arabidopsis
6	38	57.6	4092	1 DYHC_YEAST	P36022 saccharomyc
7	37.5	56.8	631	1 GIDA_BUCAP	O51879 bucherna ap
8	37	56.1	356	1 HRPW_ERWCA	Q47279 erwinia car
9	37	56.1	423	1 MK09_MOUSE	G9wtu6 mus musculu
10	37	56.1	423	1 MK09_RAT	P49186 rattus norv
11	37	56.1	620	1 Y886_METJA	O58296 methanococ
12	37	56.1	809	1 LEF_BACAN	P15917 bacillus an
13	36	54.5	213	1 RL14_RAT	O63507 rattus norv
14	36	54.5	276	1 RL2_HELPJ	G9zjr6 helicobacte
15	36	54.5	276	1 RL2_HELPJ	P56030 helicobacte
16	36	54.5	343	1 HRCA_BACSH	O69266 bacillus sp
17	36	54.5	490	1 YUE_BORBU	O51345 borrelia bu
18	36	54.5	499	1 YUAR_ECOLI	P34211 escherichia
19	36	54.5	594	1 SYFB_YEAST	P15624 saccharomyc
20	35	53.0	187	1 MCBG_ECOLI	P05530 escherichia
21	35	53.0	245	1 YCDX_ECOLI	P75914 escherichia
22	35	53.0	279	1 BMRR_BACSU	P39075 bacillus su
23	35	53.0	307	1 PYRB_HELPJ	O92m81 helicobacte
24	35	53.0	307	1 MDHP_YEAST	P32419 saccharomyc
25	35	53.0	342	1 HRCAL_LISMO	O9s5a6 listeria mo
26	35	53.0	394	1 SYTC_YEAST	P36421 saccharomyc
27	35	53.0	411	1 AK2_BACSP	O59229 bacillus sp
28	35	53.0	411	1 AK2_BACSP	O59229 bacillus sp
29	35	53.0	441	1 RGSE_HUMAN	O43566 homo sapien
30	35	53.0	443	1 EFIA_PLAFK	Q00080 daucusmodi
31	35	53.0	449	1 EF11_DAUCA	P29521 plausmodi
32	35	53.0	449	1 EF11_DAUCA	P50256 porphyra pu
33	35	53.0	456	1 EF1A_DICDI	P18624 dictyosteli

34	35	53.0	510	1 KPYC_SOLTU	P22200 solanum tub
35	35	53.0	544	1 RGSE_RAT	O08773 rattus norv
36	35	53.0	547	1 RGSE_MOUSE	P97492 mus musculu
37	35	53.0	792	1 SYL_MYCGE	P47508 mycoplasma
38	35	53.0	929	1 YDM6_SCHPO	P87137 schizosacch
39	35	53.0	2314	1 PTP2_HUMAN	P23471 homo sapien
40	35	53.0	2316	1 PTP2_RAT	O62656 rattus norv
41	34	51.5	68	1 Y002_BPHPI	P51701 bacterioph
42	34	51.5	166	1 YCGK_ALTCA	P43461 alteromonas
43	34	51.5	179	1 RL5_SYNPP6	O24701 synechococc
44	34	51.5	185	1 BCNA_CLOPE	P15935 clostridium
45	34	51.5	194	1 Y608_BUCAP	O51880 bucherna ap

ALIGNMENTS

RESULT 1

VE47_LAMB
ID VE47_LAMB STANDARD; PRT: 410 AA.
AC P03752;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1988 (Rel. 09, Last annotation update)
DE EA47 GENE PROTEIN.
OS Bacteriophage lambda.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
OC Lambda phage group.
OX NCBI_TaxID=10710;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83189071; PubMed=62211115;
RA Sanger F., Coulson A.R., Hong G.F., Hill D.F., Petersen G.B.;
RT "Nucleotide sequence of bacteriophage lambda DNA."
RL J. Mol. Biol. 162:729-773(1982).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J02459; AAA96559.1; -
DR PIR: A04377; ZEBP4L.
SQ SEQUENCE 410 AA; 48096 MW; 360E376E261EE1C1 CRC64;

Query Match 65.2%; Score 43; DB 1; Length 410;
Best Local Similarity 57.1%; Pred. No. 2.1;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 1 NERIINVSIIKLKR 14
|| :||: |||||
Db 117 NESMLNIGIKLVR 130

RESULT 2

PTGA_MYCGE
ID PTGA_MYCGE STANDARD; PRT: 908 AA.
AC P47315;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PTS SYSTEM, GLUCOSE-SPECIFIC IIABC COMPONENT (ETIABC-GLC) (GLUCOSE-
DE PERMEASE IIABC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, ABC
DE COMPONENT) (EC 2.7.1.69) (EII-GLC/EIIC-GLC).
GN PTSG OR MG069.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;

DR PROSITE; PS01129; PSI_RLU; 1.
KW Lyase.
SQ SEQUENCE 322 AA; 36588 MW; 3640CBDB83106463 CRC64;

Query Match 60.6%; Score 40; DB 1; Length 322;
Best Local Similarity 66.7%; Pred. No. 5.7;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NERIINVSIIKKL 12
II:III I:I I
DB 5 NERIINSSVKML 16

RESULT 4
NARG_BACSU
ID NARG_BACSU STANDARD; PRT; 1228 AA.
AC P42175;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE NITRATE REDUCTASE ALPHA CHAIN (EC 1.7.99.4).
NARG.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Boursier L.; Cruz Ramos H.; Danchin A.; Glaser P.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=96004574; PubMed=7557333;
RA Hoffmann T.; Group B.; Szabo A.; Hungerer C.; Jahn D.;
RT "The anaerobic life of Bacillus subtilis: cloning of the genes
encoding the respiratory nitrate reductase system.";
RL FEMS Microbiol. Lett. 131:219-225(1995).
CC !- FUNCTION: THE ALPHA CHAIN IS THE ACTUAL SITE OF NITRATE REDUCTION.
CC !- CATALYTIC ACTIVITY: NITRITE + ACCEPTOR = NITRATE + REDUCED
CC !- ACCEPTOR.
CC !- COFACTOR: MOLYBDENUM (MOLYBDOPTERIN); MAY BIND A 4FE-4S CLUSTER
CC !- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED.
CC !- SIMILARITY: BELONGS TO THE PROKARYOTIC MOLYBDOPTERIN-CONTAINING
CC OXIDOREDUCTASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL; Z49884; CAA90045.1;
DR EMBL; X91819; CAA62926.1;
DR EMBL; X85014; CAA59371.1;
DR EMBL; Z99123; CAB15736.1;
DR Subtilist; BG11081; narg.
DR InterPro; IPR001467;
DR Pfam; PF01568; Molybdop_binding; 1.
DR Pfam; PF00384; molybdoppterin; 1.
DR PROSITE; PS00551; MOLYBDOPTERIN_PROK_1; 1.
DR PROSITE; PS00490; MOLYBDOPTERIN_PROK_2; 1.
DR PROSITE; PS00932; MOLYBDOPTERIN_PROK_3; 1.
KW Nitrate assimilation; Oxidoreductase; Electron transport; Membrane;
Molybdenum; 4Fe-4S; Iron-sulfur.
FT METAL 54 54 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 58 58 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 62 62 IRON-SULFUR (4FE-4S) (BY SIMILARITY).

FT METAL 97 97 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT CONFLICT 47 47 T -> D (IN REF. 2).
SQ SEQUENCE 1228 AA; 139084 MW; D71711214393C48 CRC64;

Query Match 60.6%; Score 40; DB 1; Length 1228;
Best Local Similarity 61.5%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NERIINVSIIKKL 13
II:III I:I I
DB 1209 NORDINNVIRKLL 1221

RESULT 5
H101_ARATH
ID H101_ARATH STANDARD; PRT; 911 AA.
AC P42730;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HEAT SHOCK PROTEIN 101.
GN HSP101.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA; TISSUE=Leaf;
RX MEDLINE=95170291; PubMed=7866032;
RA Schirmer E.C.; Lindquist S.; Vierling E.;
RT "An Arabidopsis heat shock protein complements a thermotolerance
defect in yeast.";
RL Plant Cell 6:1899-1909(1994).
CC !- FUNCTION: PLAYS AN IMPORTANT ROLE IN THERMOTOLERANCE.
CC !- INDUCTION: BY HEAT SHOCK.
CC !- SIMILARITY: BELONGS TO THE CLPA/CLPB FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

EMBL; U13949; AAA67927.1;
DR InterPro; IPR001270;
DR Pfam; PF00495; clpA_B; 1.
DR PRINTS; PR00300; CLPPTHEASEA.
DR PROSITE; PS00870; CLPAB_1; 1.
DR PROSITE; PS00871; CLPAB_2; 1.
KW Chaperone; Heat shock; ATP-binding; Repeat.
FT DOMAIN 164 410 I.
FT DOMAIN 532 723 II.
FT NP_BIND 207 214 ATP (POTENTIAL).
FT NP_BIND 606 613 ATP (POTENTIAL).
SQ SEQUENCE 911 AA; 101268 MW; 0C18148A95B859BD CRC64;

Query Match 57.6%; Score 38; DB 1; Length 911;
Best Local Similarity 63.6%; Pred. No. 39;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 ERIINVSIIKKL 12
II:III I:I I
DB 63 ERVINQALKKL 73

RESULT 6
DYHC_YEAST

ID DYHC YEAST STANDARD; PRT; 4092 AA.
 AC P36022;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE DYNEIN HEAVY CHAIN, CYTOSOLIC (DYHC).
 GN DYN1 OR DHC1 OR YKR054C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9408566; PubMed=8248224;
 RA Eshel D., Urrestarazu L.A., Vissers S., Jauniaux J.-C.,
 RA van Vliet-Reedijk J.C., Planta R.J., Gibbons I.R.;
 RT "Cytoplasmic dynein is required for normal nuclear segregation in
 RT yeast."
 RL Proc. Natl. Acad. Sci. U.S.A. 90:11172-11176(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Vissers S., Urrestarazu L.A., Jauniaux J.-C.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-3457 FROM N.A.
 RX MEDLINE=94052110; PubMed=8234262;
 RA Li Y.-Y., Yeh E.-Y., Hays T., Bloom K.S.;
 RT "Disruption of mitotic spindle orientation in a yeast dynein mutant."
 RL Proc. Natl. Acad. Sci. U.S.A. 90:10096-10100(1993).
 CC -1- FUNCTION: DYNEIN HAS ATPASE ACTIVITY. CYTOPLASMIC DYNEIN ACTS AS A
 CC MOTOR FOR THE INTRACELLULAR RETROGRADE MOTILITY OF VESICLES AND
 CC ORGANELLES ALONG MICROTUBULES. REQUIRED TO MAINTAIN UNIFORM
 CC NUCLEAR DISTRIBUTION IN HYPAE. MAY PLAY AN IMPORTANT ROLE IN THE
 CC PROPER ORIENTATION OF THE MITOTIC SPINDLE INTO THE BUDDING
 CC DAUGHTER CELL YEAST. PROBABLY REQUIRED FOR NORMAL PROGRESSION OF
 CC THE CELL CYCLE.
 CC -1- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF
 CC INTERMEDIATE AND LIGHT CHAINS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. PROBABLY BINDS INDIRECTLY TO
 CC THE INNER PLASMA MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; 221877; CAA79923.1; -
 CC EMBL; L15626; AAA16055.1; -
 CC EMBL; 228279; CAA82132.1; -
 CC PIR; S38128; S38128.
 CC SGD; S0001762; DYN1.
 KW Motor protein; Microtubules; Dynein; ATP-binding; Coiled coil;
 KW Karyogamy.
 FT DOMAIN 154 175 COILED COIL (POTENTIAL).
 FT DOMAIN 486 508 COILED COIL (POTENTIAL).
 FT DOMAIN 542 566 COILED COIL (POTENTIAL).
 FT DOMAIN 932 959 COILED COIL (POTENTIAL).
 FT DOMAIN 1042 1063 COILED COIL (POTENTIAL).
 FT DOMAIN 1681 1705 COILED COIL (POTENTIAL).
 FT DOMAIN 1894 1922 MICROTUBULE-BINDING (POTENTIAL).
 FT DOMAIN 2993 3092 COILED COIL (POTENTIAL).
 FT DOMAIN 3242 3300 COILED COIL (POTENTIAL).
 FT DOMAIN 3532 3608 COILED COIL (POTENTIAL).
 FT NP_BIND 1796 1803 ATP (POTENTIAL).
 FT NP_BIND 2074 2081 ATP (POTENTIAL).
 FT NP_BIND 2418 2425 ATP (POTENTIAL).
 FT NP_BIND 2760 2767 ATP (POTENTIAL).
 FT NP_BIND 589 589
 FT CONFLICT 601 601
 Y -> C (IN REF. 3).
 V -> A (IN REF. 3).

FT CONFLICT 1364 1364 E -> A (IN REF. 3).
 FT CONFLICT 2118 2119 ML -> IV (IN REF. 1).
 SQ SEQUENCE 4092 AA; 471337 MW; 3D9DF447E8E2D6BB CRC64;
 Query Match 57.6%; Score 38; DB 1; Length 4092;
 Best Local Similarity 50.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 NERIINVSIIKKL 12
 Db 3013 NORFVNVGLEKL 3024
 I:|:|:|:|:|:|
 RESULT 7
 ID GIDA_BUCAP STANDARD; PRT; 631 AA.
 AC O51879;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GLUCOSE INHIBITED DIVISION PROTEIN A.
 GN GIDA.
 OS Buchnera aphidicola (subsp. Schizaphis graminum).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=98794;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98184963; PubMed=9516544;
 RA Clark M.A., Baumann L., Baumann P.;
 RT "Sequence analysis of a 34.7-kb DNA segment from the genome of
 RT Buchnera aphidicola (endosymbiont of aphids) containing groEL, dnaA,
 RT the atp operon, gida, and rho."
 RL Curr. Microbiol. 36:158-163(1998).
 CC -1- FUNCTION: NOT KNOWN
 CC -1- SIMILARITY: BELONGS TO THE GIDA FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF008210; AAC38117.1; -
 CC InterPro; IPR002218; -
 CC Pfam; PF01134; GIDA; 1.
 CC PROSITE; PS01280; GIDA_1; 1.
 CC PROSITE; PS01281; GIDA_2; 1.
 SQ SEQUENCE 631 AA; 70795 MW; DBBD71972F3370A3 CRC64;
 Query Match 56.8%; Score 37.5; DB 1; Length 631;
 Best Local Similarity 38.9%; Pred. No. 33;
 Matches 7; Conservative 6; Mismatches 0; Indels 5; Gaps 1;
 Qy 1 NERIINVS-----IKKLK 13
 Db 469 NEKVLNISNEKNRLKKIK 486
 I:|:|:|:|:|:|
 RESULT 8
 ID HRPN_ERWCA STANDARD; PRT; 356 AA.
 AC Q47279;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE HARPIN (HARPIN-ECC) (FRAGMENT).
 GN HRPN.
 OS Erwinia carotovora.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

CC Pectobacterium.
 CC NCBI_TaxID=554;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=71;
 CC MEDLINE=96405946; PubMed=8810071;
 CC Cui Y., Madi L., Mukherjee A., Dumenyo C.K., Chatterjee A.K.;
 CC "The RsmA-mutants of *Erwinia carotovora* subsp. *carotovora* strain
 CC Eccl1 overexpress hrpNEcc and elicit a hypersensitive reaction-like
 CC response in tobacco leaves";
 CC Mol. Plant Microbe Interact. 9:565-573(1996).
 CC -1- FUNCTION: ELICITS THE HYPERSENSITIVE RESPONSE (HR) IN THE PLANT
 CC UPON INFECTION. HARPIN ELICITS HR IN NON-HOSTS AND IS ALSO
 CC REQUIRED FOR PATHOGENICITY IN HOST PLANTS.
 CC -1- SUBCELLULAR LOCATION: SECRETED; VIA THE HRP SECRETION PATHWAY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L78834; ABA49733.1; -
 CC Hypersensitive response.
 CC DOMAIN 1 224 GLY-RICH.
 CC NON_TER 356 356
 CC SEQUENCE 356 AA; 35621 MW; 108B46B9D27F9DE4 CRC64;
 CC
 CC Query Match 56.1%; Score 37; DB 1; Length 356;
 CC Best Local Similarity 54.5%; Pred. No. 22;
 CC Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 2 ERINVSIXKL 12
 CC DB 344 DRVINGMLKLL 354
 CC
 CC RESULT 9
 CC MK09_MOUSE
 CC ID MK09_MOUSE STANDARD; PRT; 423 AA.
 CC AC Q9WTU6; Q9WTU4; Q9WTU5;
 CC DT 01-OCT-2000 (Rel. 40, Created)
 CC DT 01-OCT-2000 (Rel. 40, Last sequence update)
 CC DT 01-OCT-2000 (Rel. 40, Last annotation update)
 CC DE MITOGEN-ACTIVATED PROTEIN KINASE 9 (EC 2.7.1.-) (STRESS-ACTIVATED
 CC PROTEIN KINASE JNK2) (C-JUN N-TERMINAL KINASE 2).
 CC GN MAPK9 OR PRKM9 OR JNK2.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 CC OX NCBI_TaxID=10090;
 CC [1]
 CC SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 CC Barrett T., Davis R.J.;
 CC "JNK2 is required for the development of Th1 cells";
 CC Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC [2]
 CC SEQUENCE FROM N.A. (ISOFORM ALPHA-1).
 CC TISSUE=Brain;
 CC Ito M., Yoshioka K., Akechi M., Yamashita S., Takamatsu N.,
 CC Sugiyama K., Hibi M., Nakabeppu Y., Shiba T., Yamamoto K.I.;
 CC "JSAPl, a novel jun N-terminal protein kinase (JNK)-binding protein
 CC that functions as a scaffold factor in the JNK signaling pathway";
 CC Mol. Cell. Biol. 19:7539-7548(1999).
 CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN THE SIGNAL TRANSDUCTION PATHWAY
 CC INITIATED BY PROINFLAMMATORY CYTOKINES AND UV RADIATION.
 CC -1- ENZYME REGULATION: ACTIVATED BY THREONINE AND TYROSINE
 CC PHOSPHORYLATION (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; ALPHA-1, ALPHA-2 (SHOWN HERE),
 CC BETA-1 AND BETA-2; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- PTM: AUTOPHOSPHORYLATED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF052469; AAD22579.1; -
 CC EMBL: AF052466; AAD22576.1; -
 CC EMBL: AF052468; AAD22578.1; -
 CC EMBL: AB005664; BAA85876.1; -
 CC MGD: MGI:1346862; Mapk9.
 CC HSP: P47811; IP38.
 CC InterPro: IPR000719; -
 CC InterPro: IPR002290; -
 CC PROSITE: PS01351; MAPK; 1.
 CC PROSITE: PS00107; PROTEIN_KINASE_DOM; 1.
 CC PROSITE: PS00011; PROTEIN_KINASE_ST; 1.
 CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 CC Transferase; Serine/threonine-protein kinase; ATP-binding;
 CC Phosphorylation; Alternative splicing.
 CC DOMAIN 26 321 PROTEIN KINASE.
 CC NP_BIND 32 40 ATP (BY SIMILARITY).
 CC BINDING 55 55 ATP (BY SIMILARITY).
 CC ACT_SITE 151 151 BY SIMILARITY.
 CC MOD_RES 183 183 PHOSPHORYLATION (ACTIVATES THE KINASE).
 CC MOD_RES 185 185 PHOSPHORYLATION (ACTIVATES THE KINASE).
 CC VARSPLIC 216 230 AEMVLHKVLPGRDY -> GELVKGCVIFQGTDH (IN
 CC ISOFORM BETA-1 AND ISOFORM BETA-2).
 CC VARSPLIC 377 423 DAAVSKATPQSSSINDISMTEHTLIADTSSLDASTG
 CC PLEGR -> AQMQQ (IN ISOFORM ALPHA-1 AND
 CC ISOFORM BETA-1).
 CC SQ SEQUENCE 423 AA; 48189 MW; 0E759B486ABCE20D CRC64;
 CC
 CC Query Match 56.1%; Score 37; DB 1; Length 423;
 CC Best Local Similarity 70.0%; Pred. No. 27;
 CC Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 5 INVSIIKLKR 14
 CC DB 50 INAVKKLSR 59
 CC
 CC RESULT 10
 CC MK09_RAT
 CC ID MK09_RAT STANDARD; PRT; 423 AA.
 CC AC P49186;
 CC DT 01-FEB-1996 (Rel. 33, Created)
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
 CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
 CC DE MITOGEN-ACTIVATED PROTEIN KINASE 9 (EC 2.7.1.-) (STRESS-ACTIVATED
 CC PROTEIN KINASE JNK2) (C-JUN N-TERMINAL KINASE 2) (SAPK-ALPHA) (P54-
 CC ALPHA).
 CC GN MAPK9 OR PRKM9 OR JNK2.
 CC OS Rattus norvegicus (Rat).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC OX NCBI_TaxID=10116;
 CC [1]
 CC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 CC TISSUE=Brain;
 CC MEDLINE=94232348; PubMed=8177321;
 CC Kyriakis J.M., Banerjee P., Nikolakaki E., Dai T., Rubie E.A.,
 CC Ahmad M.F., Avruch J., Woodgett J.R.;
 CC "The stress-activated protein kinase subfamily of c-Jun kinases";
 CC Nature 369:156-160(1994).
 CC -1- FUNCTION: PROBABLY PLAYS A ROLE IN THE SIGNAL TRANSDUCTION PATHWAY

CC INITIATED BY PROINFLAMMATORY CYTOKINES, UV RADIATION AND HEAT
 CC SHOCK. BINDS TO THE N-TERMINAL ACTIVATION DOMAIN OF C-JUN AND
 CC PHOSPHORYLATES THE REGULATORY SITES SER-63 AND SER-73.
 CC -1- ENZYME REGULATION: ACTIVATED BY THREONINE AND TYROSINE
 CC PHOSPHORYLATION.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: ALPHA-2 (SHOWN HERE) AND ALPHA-
 CC 1: ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L27112; AAA42109.1; -
 CC EMBL: L27111; AAA42108.1; -
 CC HSP60. Q16539; IWC.
 CC InterPro: IPR000719; -
 CC InterPro: IPR002290; -
 CC Pfam: PF00069; pkinase; 1.
 CC PROSITE: PS01351; MAPK; 1.
 CC PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 CC PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 CC Transferase: Serine/threonine-protein kinase; ATP-binding;
 CC Phosphorylation; Alternative splicing.
 CC DOMAIN 26 321
 CC NP_BIND 32 40
 CC BINDING 55 55
 CC ACT_SITE 151 151
 CC MOD_RES 183 183
 CC MOD_RES 185 185
 CC VARSPLIC 216 230
 CC ISOFORM ALPHA-1).
 CC SEQUENCE 423 AA; 48017 MW; EE549B9FAF12F421 CRC64;
 CC -----
 CC Query Match 56.1%; Score 37; DB 1; Length 423;
 CC Best Local Similarity 70.0%; Pred. No. 27;
 CC Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC QY 5 INVSIKKLKR 14
 CC III:IIII I
 CC Db 50 INAVKKLSR 59
 CC -----
 CC RESULT 11
 CC Y886_METJA STANDARD; PRT; 620 AA.
 CC AC Q58296;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DE PUTATIVE MOLYBDOPTERIN BIOSYNTHESIS PROTEIN MJ0886.
 CC GN MJ0886.
 CC OS Methanococcus jannaschii.
 CC OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 CC OC Methanococcus.
 CC OX NCBI_TaxID=2190;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 CC RA MEDLINE=9637999; PubMed=8688087;
 CC RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 CC RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 CC RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 CC RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 CC RA Scott J.B., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 CC RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,

RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RA jannaschii.";
 RA Science 273:1058-1073(1996).
 CC -1- PATHWAY: MOLYBDENUM COFACTOR BIOSYNTHESIS (BY SIMILARITY).
 CC -1- SIMILARITY: STRONG, IN THE N-TERMINAL, TO BACTERIAL MOEA PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U67532; AAB98890.1; -
 CC TIGR: MJ0886;
 CC InterPro: IPR001453; -
 CC Pfam: PF00994; MOCF_biosynth; 1.
 CC PROSITE: PS01079; MOCF_BIOSYNTHESIS_2; 1.
 CC KW Hypothetical protein; Molybdenum cofactor biosynthesis.
 CC SEQUENCE 620 AA; 68896 MW; 9B94AAA7CA6DFOAB CRC64;
 CC -----
 CC Query Match 56.1%; Score 37; DB 1; Length 620;
 CC Best Local Similarity 80.0%; Pred. No. 40;
 CC Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 CC -----
 CC QY 4 IINVSIIKLK 13
 CC III:IIII I
 CC Db 16 IINESLKLK 25
 CC -----
 CC RESULT 12
 CC LEF_BACAN STANDARD; PRT; 809 AA.
 CC ID LEF_BACAN
 CC AC P15917;
 CC DT 01-APR-1990 (Rel. 14, Created)
 CC DT 01-APR-1990 (Rel. 14, Last sequence update)
 CC DT 15-JUL-1998 (Rel. 36, Last annotation update)
 CC DE LETHAL FACTOR PRECURSOR (EC 3.4.24.-) (LF).
 CC GN LEF.
 CC OS Bacillus anthracis.
 CC OG Plasmid pXOI.
 CC OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC OC Bacillus/Staphylococcus group; Bacillus.
 CC OX NCBI_TaxID=1392;
 CC RN [1]
 CC RP SEQUENCE FROM N.A., AND SEQUENCE OF 34-49.
 CC RX MEDLINE=90034185; PubMed=2509294;
 CC RA Bragg T.S., Robertson D.L.;
 CC RT "Nucleotide sequence and analysis of the lethal factor gene (lef)
 CC RT from Bacillus anthracis.";
 CC RL Gene 81:45-54(1989).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RA Lowe J.;
 CC RL Submitted (APR-1990) to the EMBL/GenBank/DBJ databases.
 CC RN [3]
 CC RP ZINC-BINDING.
 CC RX MEDLINE=95154669; PubMed=7851740;
 CC RA Kochi S.K., Schiavo G., Mock M., Montecucco C.;
 CC RT "Zinc content of the Bacillus anthracis lethal factor.";
 CC RL FEMS Microbiol. Lett. 124:343-348(1994).
 CC CC -1- FUNCTION: ONE OF THE THREE PROTEINS COMPOSING THE ANTHRAX TOXIN,
 CC AGENT WHICH INFECTS MANY MAMMALIAN SPECIES AND THAT MAY CAUSE
 CC DEATH. LF IS THOUGHT TO BE A LETHAL FACTOR THAT, WHEN ASSOCIATED
 CC WITH PA, CAUSES DEATH. LF IS NOT TOXIC BY ITSELF. PA IS THOUGHT TO
 CC BIND TO RECEPTORS OF SENSITIVE EUKARYOTIC CELLS, THEREBY
 CC FACILITATING THE INTERNALIZATION OF LF OR EF.
 CC -1- SUBUNIT: SECRETED ANTHRAX TOXIN IS COMPOSED OF THREE DISTINCT
 CC PROTEINS, A PROTECTIVE ANTIGEN (PA), A LETHAL FACTOR (LF) AND AN


```

QY 1 NERIINVSIIKKLR 14
   ||| ||| |
Db 198 NEDFINSIGRAGR 211

RESULT 15
RL2_HELPY
ID RL2_HELPY STANDARD; PRT; 276 AA.
AC PS6030;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L2.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RA MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
pylori.";
RL Nature 388:539-547(1997).
CC -!- FUNCTION: THIS PROTEIN IS A PRIMARY 23S rRNA-BINDING PROTEIN. IT
CC HAS PEPTIDYLTRANSFERASE ACTIVITY (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE L2P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000633; AAD08355.1; -
CC TIGR; HP1316; -
CC InterPro: IPR002171; -
CC Pfam; PF00181; Ribosomal_L2; 1.
CC PROSITE; PS00467; RIBOSOMAL_L2; 1.
KW Ribosomal protein; rRNA-binding.
SQ SEQUENCE 276 AA; 30271 MW; 5DCA6F36726C9A35 CRC64;

Query Match 54.5%; Score 36; DB 1; Length 276;
Best local Similarity 64.3%; Pred. No. 26;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 NERIINVSIIKKLR 14
   ||| ||| |
Db 198 NEDFINSIGRAGR 211

Search completed: May 23, 2001, 11:20:27
Job time: 525 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:19:46 ; Search time 118.04 Seconds
(Without alignments)
13.901 Million cell updates/sec

Title: US-09-522-217-2_COPY_92_105

Perfect score: 66

Sequence: 1 NERIINVSIIKKLR 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: SPTREMBL_15.*
- 2: sp-archaea.*
- 3: sp-bacteria.*
- 4: sp-fungi.*
- 5: sp-human.*
- 6: sp-invertebrate.*
- 7: sp-mammal.*
- 8: sp-mhc.*
- 9: sp-organelle.*
- 10: sp-phage.*
- 11: sp-plant.*
- 12: sp-rodent.*
- 13: sp-unclassified.*
- 14: sp-virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	ID	Description
1	43	65.2	620	2	Q33997 chromatin
2	41	62.1	284	5	Q9NIY1 trypanodendro
3	41	62.1	615	10	Q9SMQ3 arabidopsis
4	40	60.6	276	2	Q9KJ70 vibrio chol
5	40	60.6	292	5	Q16616 caenorhabdi
6	40	60.6	385	5	Q9NA54 caenorhabdi
7	40	60.6	489	10	Q9LYL3 arabidopsis
8	40	60.6	602	2	Q9KGU3 leptospira
9	40	60.6	630	5	Q27733 plasmodium
10	39	59.1	210	2	Q9PNL3 campylobact
11	39	59.1	223	2	Q31432 bacillus su
12	39	59.1	417	2	Q67158 aquifex aeo
13	39	59.1	572	1	Q59663 pyrodictum
14	39	59.1	576	11	Q62970 rattus norv
15	39	59.1	788	3	Q07381 saccharomyc
16	38	57.6	208	1	Q30253 archaeoglob
17	38	57.6	247	8	Q9MUR0 mesostigma
18	38	57.6	292	5	Q16679 caenorhabdi
19	38	57.6	333	2	Q54206 streptomyce

20	38	57.6	702	3	Q04924
21	38	57.6	911	10	Q9LE57
22	38	57.6	912	10	Q9S822
23	38	57.6	1245	5	Q96195
24	37	56.1	171	2	Q9PFJ4
25	37	56.1	283	1	Q26701
26	37	56.1	334	8	Q33796
27	37	56.1	364	5	Q02423
28	37	56.1	364	5	Q02449
29	37	56.1	381	11	Q9WTU4
30	37	56.1	381	11	Q9WTU5
31	37	56.1	382	13	P79996
32	37	56.1	391	5	Q17791
33	37	56.1	423	11	Q9WTU6
34	37	56.1	482	11	P70320
35	37	56.1	549	4	Q9NYH2
36	37	56.1	717	11	O55047
37	37	56.1	718	4	Q9UKI8
38	37	56.1	719	4	Q9Y4F7
39	37	56.1	741	2	O24875
40	37	56.1	742	2	Q92N31
41	37	56.1	749	4	Q9UKI7
42	37	56.1	787	4	Q9Y4F6
43	37	56.1	801	4	Q14150
44	37	56.1	883	5	Q07995
45	37	56.1	1089	5	Q9VHC4

ALIGNMENTS

RESULT 1	
O33997	
ID O33997	PRELIMINARY; PRT; 620 AA.
AC O33997	
DT 01-JAN-1998	(TREMBLrel. 05, Created)
DT 01-JAN-1998	(TREMBLrel. 05, Last sequence update)
DT 01-MAY-2000	(TREMBLrel. 13, Last annotation update)
DE	ADENYLYLSULFATE REDUCTASE ALPHA SUBUNIT.
GN	APRA.
OS	Chromatium vinosum.
OC	Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;
OC	Allochrochromatium.
OX	NCBI_TaxID=1049;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=D DSM 180;
RX	MEDLINE=7453474; PubMed=9308173;
RA	Hipp W.M., Pott A.S., Thum-Schmitz N., Faath I., Dahl C.,
RA	Trueper H.G.;
RT	"Towards the phylogeny of APS reductases and sirohaem sulfite
RT	reductases in sulfate-reducing and sulfur-oxidizing prokaryotes.";
RL	Microbiology 143:0-0(0).
DR	EMBL; U84759; AAC23621.1;
DR	INTERPRO; IPR000464;
DR	PFAM; PF00890; FAD-binding_2; 1.
SQ	SEQUENCE 620 AA; 69543 MW; 9FCFD07B77DA34A8 CRC64;

Query Match 65.2%; Score 43; DB 2; Length 620;
Best Local Similarity 46.2%; Pred. No. 41;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 NERIINVSIIKKLR 13
|| ::| :||| :|

Db 518 NENLLNIGLKKMK 530

RESULT 2

Q9NIY1 PRELIMINARY; PRT; 284 AA.
ID Q9NIY1
AC Q9NIY1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE ELONGATION FACTOR 1 ALPHA (FRAGMENT).
 GN EF-1A.
 OS Trypoderon lineatum.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
 OC Cucujiformia; Scolytidae; Trypoderon.
 OX NCBI_TaxID=105211;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Normark B.B., Jordal B.H., Farrell B.D.;
 RT "Origin of a haplodiploid beetle lineage";
 RL Proc. R. Soc. Lond., B, Biol. Sci. 266:2253-2259(1999).
 DR EMBL; AF186682; AAF33736.1; -;
 FT NON_TER 284 284
 FT NON_TER 1 1
 SQ SEQUENCE 284 AA; 30995 MW; C2C4A6C63C01C79B CRC64;

Query Match 62.1%; Score 41; DB 5; Length 284;
 Best Local Similarity 63.6%; Pred. No. 42;
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 4 IINVSIIKKLR 14
 Db 253 VINSVRELRR 263

RESULT 3
 Q9SMQ3 PRELIMINARY; PRT; 615 AA.
 AC Q9SMQ3;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE HYPOTHETICAL 70.0 KDA PROTEIN.
 GN T5017.190 OR AT4G40020.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
 RA Mayer K.F.X., Lemcke K., Schueller C.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL035708; CAB38915.1; -;
 DR EMBL; AL161596; CAB80665.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 615 AA; 70024 MW; ED8B0287B2F3E456 CRC64;

Query Match 62.1%; Score 41; DB 10; Length 615;
 Best Local Similarity 57.1%; Pred. No. 87;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 NERIINVSIIKKLR 14
 Db 350 NEAVANDNIKKLR 363

RESULT 4

Q9KLJ0 PRELIMINARY; PRT; 276 AA.
 ID Q9KLJ0;
 AC Q9KLJ0;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE LIPASE-RELATED PROTEIN.
 DE VCA0754.
 GN Vibrio cholerae.
 OS Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae";
 RL Nature 406:477-483(2000).
 DR EMBL; AE004404; AAF96652.1; -;
 DR TIGR; VCA0754; -;
 SQ SEQUENCE 276 AA; 32570 MW; E3B9194B67B12C11 CRC64;

Query Match 60.6%; Score 40; DB 2; Length 276;
 Best Local Similarity 66.7%; Pred. No. 59;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RIINVSIIKKLR 14
 Db 9 RIIDVTVKPLKR 20

RESULT 5

O16616 PRELIMINARY; PRT; 292 AA.
 ID O16616;
 AC O16616;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE B0281.8 PROTEIN.
 GN B0281.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peleoderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,
 RA Smaildon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans";
 RL Nature 368:32-38(1994).
 RN [2]

```

RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Pauley A., Scheet P.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RC -1- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
DR EMBL; AF016666; AAB66086.1; -.
DR INTERPRO; IPR001841; -.
DR PFAM; PF00097; zf-C3HC4; 1.
DR PROSITE; PS00518; ZINC_FINGER_C3HC4; 1.
KW Zinc-finger.
SQ
SEQUENCE 292 AA; 33858 MW; DF40855A856D49B7 CRC64;

```

Query Match 60.6%; Score 40; DB 5; Length 292;
Best Local Similarity 38.5%; Pred. No. 62;
Matches 5; Conservative 7; Mismatches 1; Indels

QY 1 NERIINVSIIKLLK 13
:|::|: :||:|
Db 176 SEOLLNIKVKKIK 188

RESULT	6		
Q9NA54			
ID	Q9NA54	PRELIMINARY;	385 AA.
AC	Q9NA54;		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)		
DE	Y73F8A.18 PROTEIN.		
DE	Y73F8A.18		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhaditida; Rhaditidae; Caenorhabditis.		
OC	Rhabditidae; Peleoderinae; Caenorhabditis.		
OX	NCBI_TaxID=6239;		
RN	[1]		
RN	SEQUENCE FROM N.A.		
RP	Matthews L.;		
RA			
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	none;		
RT	"Genome sequence of the nematode C.elegans: A platform for investigating biology."		
RL	Science 282:2012-2018(1998).		
RL	EMBL; ALI32862; CAB60543.1; .		
DR	SEQUENCE 385 AA; 42966 MW; 75F1D8F24DFB83C CRC64;		

Query Match 60.6%; Score 40; DB 5; Length 385;
Best Local Similarity 50.0%; Pred. No. 81;
Matches 7; Conservative 3; Mismatches 4; Indels

Qy 1 NERIINVSIIKKLR 14
| : : : | : : :
Db 282 NPALLNISEPKLR 298

RESULT	7	
Q9LYL3		
ID	Q9LYL3	PRELIMINARY;
AC	Q9LYL3	PRT; 489 AA.
DT	01-OCT-2000	{TremBrel. 15, Created}
DT	01-OCT-2000	{TremBrel. 15, Last sequence update}
DT	01-OCT-2000	{TremBrel. 15, Last annotation update}
DE	CYSTEINE-TRNA LIGASE.	
FI	F18021.260.	
OS	Arabidopsis thaliana (Mouse-ear cress).	

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II
 OC Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 [1]
 RN SEQUENCE FROM N.A.
 RA Benes V., Wurmbach E., Drzonek H., Ansorge W., Mewes H.W., Rudd S.
 RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RP EU Arabidopsis sequencing project;
 RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL163763; CAB87429.1; -;
 KW Ligase.
 SO SEQUENCE 489 AA: 55350 MW: EBABA9A64E0DC5DE CRC64

Query Match 60.6%; Score 40; DB 10; Length 489;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 8; Conservative 2; Mismatches 2; Indels

Qy 3 RIINVSIKLKR 14
Db 343 KFINVSISKLKK 354
: ||||| |||:

RESULT	8	
Q9KGU3		
ID	Q9KGU3	PRELIMINARY; PRT; 602 AA.
AC	Q9KGU3;	
DT	01-OCT-2000	(TrEMBLrel. 15, Created)
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)
DT	01-OCT-2000	(TrEMBLrel. 15, Last annotation update)
DE	PENICILLIN-BINDING PROTEIN 3.	
DE	PPBP.	
GN	Leptosira interrogans.	
OS	Leptosira interrogans.	
OC	Bacteria; Spirochaetales; Leptospiraceae; Leptosira.	
OX	NCBI_TaxId=173;	
RI	[1]	
RN	SEQUENCE FROM N.A.	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=R211;	
RA	Brenot A., Trott D., Saint Girons I., Zuerner R.;	
RT	"penicillin-binding proteins in leptosira interrogans.";	
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF282907; AAF87305.1;	
DR	EMBL; AF282907; AAF87305.1;	
SQ	SEQUENCE	602 AA; 67208 MW; 1A7186B2C7E1670D CRC64;

Query Match 50.6%; Score 40; DB 2; Length 602;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 4; Mismatches 2; Indels

Qy 2 ERIINVSIIKKLR 14
|| :|||:|: :|
Db 528 ERTLNVSLKRFOR 540

RESULT	9	
Q27733		
ID	Q27733	PRELIMINARY;
AC	Q27733	PRT; 630 AA.
DT	01-NOV-1996	(TrEMBLrel. 01, Created)
DT	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)
DT	01-MAY-2000	(TrEMBLrel. 13, Last annotation update)
DE	DELTA-AMINOLEVULINIC ACID SYNTHETASE (EC 2.3.1.37)	
DE	(5-AMINOLEVULINIC ACID SYNTHASE) (DELTA-AMINOLEVULINATE SYNTHASE)	
DE	(DELTA-ALA SYNTHETASE) (ALAS).	
OS	Plasmodium falciparum.	
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
OX	NCBI_TaxID=5833;	
RN	[1]	
RP	SEQUENCE FROM N.A.	

RX MEDLINE=96258568; PubMed=8992326;
RA Wilson C.M., Smith A.B., Baylon R.V.;
RT "Characterization of the delta-aminolevulinat synthase gene homologue
in P. falciparum.";
RL Mol. Biochem. Parasitol. 75:271-276(1995).
CC -1- CATALYTIC ACTIVITY: SUCCINYL-COA + GLYCINE = 5-AMINOLEVULINATE +
COA + CO(2).
CC -1- COFACTOR: PYRIDOXAL-PHOSPHATE.
DR ENBL; L46348; AAC37294.1; -.
DR INTERPRO; IPR001917; -.
DR PFAM; PF00222; aminotran_2; 1.
DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; UNKNOWN_1.
KW Transferrase; Acyltransferase.
SQ SEQUENCE 630 AA; 73546 MW; E97B61D8E97B0645 CRC64;

Query Match 60.6%; Score 40; DB 5; Length 630;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 NERIINVSIIKKLR 14
Db 274 NEKIIIEVGIELTKK 287

RESULT 10
Q9PNL3 PRELIMINARY; PRT; 210 AA.
AC Q9PNL3
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE THIAMIN-PHOSPHATE PYROPHOSPHORYLASE (EC 2.5.1.3).
GN THIE.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth F., Davies R.M., Feltham T., Holtroyd S.,
RA Jagsels K., Karlyshev A., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., Van Vliet A.H.M.,
RA Whitehead S., Barrett B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
DR ENBL; AL139077; CAB73336.1; -.
SQ SEQUENCE 210 AA; 23001 MW; 753C83D43B0EB44A CRC64;

Query Match 59.1%; Score 39; DB 2; Length 210;
Best Local Similarity 50.0%; Pred. No. 67;
Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
QY 1 NERIINVSIIKKLR 14
Db 103 DEKIIGLSLKKLEQ 116

RESULT 11
Q31432 PRELIMINARY; PRT; 223 AA.
AC Q31432
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE YBDJ PROTEIN.
GN YBDJ
OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschl C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Enrian K.D., Errington J., Fabret C., Ferreri E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takenaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tostato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus
subtilis.";
RL Nature 390:249-256(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RA Submitted (NOV-1997) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Haga K., Liu H., Yasumoto K., Takahashi H., Yoshikawa H.;
RT "Sequence analysis of the 70kb region between 17 and 23 degree of the
Bacillus subtilis chromosome.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBSJ databases.
DR ENBL; Z99105; CAB11994.1; -.
DR EMBL; AB006424; BAA33098.1; -.
DR HSSP; P08402; 1B00.
DR INTERPRO; IPR001789; -.
DR INTERPRO; IPR001867; -.
DR PFAM; PF00072; response_reg; 1.
DR PFAM; PF00486; trans_reg_C; 1.
SQ SEQUENCE 223 AA; 26173 MW; 9DB646AD90675A86 CRC64;

Query Match 59.1%; Score 39; DB 2; Length 223;
Best Local Similarity 53.8%; Pred. No. 71;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 NERIINVSIIKKLR 13
Db 184 NPNIVNHIKKIR 196

RESULT 12
O67158 PRELIMINARY; PRT; 417 AA.
ID O67158
AC O67158;

DT 01-AUG-1998 (TReMBLrel. 07, Created)
 DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
 DE HYPOTHETICAL 49.1 KDA PROTEIN.
 GN AQ_1059.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus";
 RL Nature 392:353-358(1998).
 DR EMBL; AF000721; RAC07128.1;
 KW Hypothetical protein.
 SQ SEQUENCE 417 AA; 49062 MW; F889585802AE13C CRC64;

Query Match 59.1%; Score 39; DB 2; Length 417;
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 2 ERIINVSIRKLK 13
 |:::|:::|
 Db 102 EKLIDISVKELK 113

RESULT 13
 O59663
 ID O59663 PRELIMINARY; PRT; 572 AA.
 AC O59663;
 DT 01-AUG-1998 (TReMBLrel. 07, Created)
 DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
 DE THSB PROTEIN.
 GN THSB.
 OS Pyrodicticum occultum.
 OC Archaea; Crenarchaeota; Desulfurococcales; Pyrodictiaceae;
 OC Pyrodictium.
 OX NCBI_TaxID=2309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PL19;
 RA Frey G.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PL19;
 RA Stetter K.O.;
 RT "ultrathin mycelia-forming organisms from submarine volcanic areas
 RT having an optimum growth temperature of 105 degrees Celsius.";
 RL Nature 300:258-260(1982).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PL19;
 RA Stetter K.O., Koenig H., Stackbrandt E.;
 RT "Pyrodicticum gen. nov., a new genus of submarine, disc-shaped sulfur-
 RT reducing archaeobacteria growing optimally at 105 degC";
 RL Syst. Appl. Microbiol. 4:535-551(1983).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PL19;
 RX MEDLINE=91266899; PubMed=1828761;
 RA Phipps B.M., Hoffmann A., Stetter K.O., Baumeister W.;
 RT "A novel ATPase complex selectively accumulated upon heat shock is a
 RT major cellular component of thermophilic archaeobacteria";
 RL EMBO J. 10:1711-1722(1991).
 RN [5]

RP SEQUENCE FROM N.A.
 RC STRAIN=PL19;
 RA Phipps B.M., Typke D., Hegerl R., Volker S., Hoffmann A.,
 RA Stetter K.O., Baumeister W.;
 RT "Structure of a molecular chaperone from a thermophilic
 RT archaeobacterium";
 RL Nature 361:475-477(1993).
 DR EMBL; AJ006550; CAA07096.1;
 DR HSSP; P48425; 1A6D.
 DR INTERPRO; IPR002194;
 DR INTERPRO; IPR002423;
 DR PFAM; PF00118; cpn60_TCP1_1.
 DR PRINTS; PR00304; TCOMPLEXTCP1.
 DR PROSITE; PS00750; TCP1_1; 1.
 DR PROSITE; PS00751; TCP1_2; 1.
 DR PROSITE; PS00995; TCP1_3; 1.
 SQ SEQUENCE 572 AA; 62763 MW; C2B241B7D71AD45D CRC64;

Query Match 59.1%; Score 39; DB 1; Length 572;
 Best Local Similarity 46.2%; Pred. No. 1.7e+02;
 Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 2 ERIINVSIRKLK 14
 |:::|:::|
 Db 295 EKYNVAVERMKR 307

RESULT 14
 Q62970
 ID Q62970 PRELIMINARY; PRT; 576 AA.
 AC Q62970;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)
 DE APOLIPOPROTEIN B (FRAGMENT).
 GN APOB.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Shimmin L.C.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U53873; AAA98613.1;
 KW Lipoprotein.
 FT NON_TER 1 576
 FT NON_TER 576 576
 SQ SEQUENCE 576 AA; 66762 MW; FF649CCAA69AA07A CRC64;

Query Match 59.1%; Score 39; DB 11; Length 576;
 Best Local Similarity 63.6%; Pred. No. 1.7e+02;
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 RIINVSIRKLK 13
 |:::|:::|
 Db 198 RLIDMSVKLK 208

RESULT 15
 Q07381
 ID Q07381 PRELIMINARY; PRT; 788 AA.
 AC Q07381;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
 DE CHROMOSOME IV READING FRAME ORF YDL060W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Bloecker H., Brandt P.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z74108; CAA98623.1; -
DR INTERPRO; IPR000985; -
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; UNKNOWN_1.
SQ SEQUENCE 788 AA; 90747 MW; 71BA5461E8A858C7 CRC64;

Query Match 59.1%; Score 39; DB 3; Length 788;
Best Local Similarity 61.5%; Pred. No. 2.4e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NERIINVSIIKKLK 13
Db 120 SKRIFNVHKKFK 132

Search completed: May 23, 2001, 11:19:47
Job time: 600 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:11:38 ; Search time 108.07 Seconds

(without alignments)
7.405 Million cell updates/sec

Title: US-09-522-217-2_COPY_135_148

Perfect score: 67

Sequence: 1 EFLERFKSLQKMI 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_0401.*
1: /SIDS6/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS6/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS6/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS6/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS6/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS6/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS6/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS6/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS6/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS6/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS6/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS6/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS6/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS6/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS6/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS6/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS6/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS6/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS6/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS6/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS6/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS6/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	67	100.0	32	B18626	Antigenic peptide
2	67	100.0	162	21	A human zalphall 1
3	67	100.0	519	21	Amino acid sequenc
4	54	80.6	146	21	A mouse zalphall 1
5	54	80.6	510	21	Amino acid sequenc
6	42	62.7	133	18	Feline interleukin
7	40	59.7	277	21	Arabidopsis thalia
8	40	59.7	304	21	Arabidopsis thalia
9	40	59.7	305	21	Arabidopsis thalia
10	39	58.2	23	17	Lytic peptide used
11	39	58.2	23	17	Synthetic lytic pe

12	39	58.2	27	17	R92429	Lytic peptide used
13	39	58.2	27	17	R92430	Lytic peptide used
14	39	58.2	27	17	R89986	Synthetic lytic pe
15	39	58.2	27	17	R89987	Synthetic lytic pe
16	39	58.2	108	21	Y58205	Canine mature inte
17	39	58.2	132	21	Y58203	Canine interleukin
18	39	58.2	290	21	G23634	Arabidopsis thalia
19	39	58.2	307	21	G23633	Arabidopsis thalia
20	39	58.2	310	21	G23632	Arabidopsis thalia
21	39	58.2	1817	21	B18255	Plasmodium falci
22	38	56.7	271	20	Y15227	Human receptor pro
23	38	56.7	271	21	B28205	Novel human protei
24	38	56.7	276	21	B42611	Human ORFX ORF2375
25	38	56.7	282	21	B57116	Human prostate can
26	38	56.7	642	21	B42942	Human ORFX ORF2706
27	37	55.2	389	21	B42293	Human ORFX ORF2057
28	37	55.2	600	20	W82660	Caulliflower L-gala
29	37	55.2	615	18	W20875	H. pylori cytoplas
30	37	55.2	1008	21	G32188	Arabidopsis thalia
31	37	55.2	1076	21	G32187	Arabidopsis thalia
32	37	55.2	1288	21	G32186	Arabidopsis thalia
33	36	53.7	22	20	Y19069	Lecithin:cholester
34	36	53.7	22	20	Y18956	Lecithin:cholester
35	36	53.7	22	20	Y18950	Lecithin:cholester
36	36	53.7	22	20	Y18815	Lecithin:cholester
37	36	53.7	22	20	Y18702	Lecithin:cholester
38	36	53.7	22	20	Y18696	Lecithin:cholester
39	36	53.7	22	20	Y18552	Lecithin:cholester
40	36	53.7	22	20	Y18433	Lecithin:cholester
41	36	53.7	22	20	Y18439	Lecithin:cholester
42	36	53.7	22	20	Y19204	Lecithin:cholester
43	36	53.7	22	20	Y19210	Lecithin:cholester
44	36	53.7	22	20	Y19323	Lecithin:cholester
45	36	53.7	152	20	Y35363	Chlamydia pneumoni

ALIGNMENTS

RESULT 1
B18626
ID B18626 standard; Peptide; 32 AA.
XX B18626;
XX
XX 22-JAN-2001 (first entry)
XX
XX Antigenic peptide derived from a human zalphall ligand polypeptide.
XX
XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.
XX
XX Homo sapiens.
XX
XX WO200053761-A2.
XX
XX 14-SEP-2000.
XX
XX 09-MAR-2000; 2000WO-US060607.
XX
XX 09-MAR-1999; 99US-0264908.
PR 11-MAR-1999; 99US-0265992.
PR 01-JUL-1999; 99US-0142013.
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI Johnston JV, Nelson AJ, Dillion SR, Hammond AK;
XX WPI; 2000-565600/52.
XX New human cytokine, designated zalphall ligand, useful for stimulating
PT the proliferation and/or development of haematopoietic cells in vitro

PT and in vivo, and for treating tumourigenesis -
 PS Example 34; Page 227; 256pp; English.

XX The present sequence was used to raise antibodies, and is derived from
 CC a human zalphall ligand polypeptide. Zalphall ligand is a cytokine.
 CC The zalphall ligand is useful for stimulating the proliferation and
 CC development of haematopoietic cells in vitro and in vivo. Zalphall
 CC ligand polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.

XX SQ Sequence 32 AA;

Query Match 100.0%; Score 67; DB 21; Length 32;
 Best Local Similarity 100.0%; Pred. No. 9.7e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFLERFKSLQKMI 14
 Db 14 eflerfksllqkmi 27

RESULT 2

B18623
 ID B18623 standard; Protein; 162 AA.

AC B18623;

XX 22-JAN-2001 (first entry)

XX A human zalphall ligand polypeptide.

XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.

XX Homo sapiens.

XX WO200053761-A2.

XX 14-SEP-2000.

XX 09-MAR-2000; 2000WO-US06067.

XX 09-MAR-1999; 99US-0264908.

XX 11-MAR-1999; 99US-0265992.

XX 01-JUL-1999; 99US-0142013.

XX (Zymo) ZYMOGENETICS INC.

XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;

XX WPI; 2000-565600/52.

XX N-PSDB; A75552.

XX New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -
 XX Disclosure; Page 205-206; 256pp; English.

XX

CC The present sequence represents a human zalphall ligand polypeptide,
 CC which is a cytokine. The zalphall ligand is useful for stimulating the
 CC proliferation and development of haematopoietic cells in vitro and in
 CC vivo. Zalphall ligand polynucleotides can be used as primers or probes
 CC for cloning the zalphall gene. The zalphall ligand is useful for
 CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
 CC used for treating leukaemias and lymphomas. Antagonists against zalphall
 CC ligand are useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.

XX SQ Sequence 162 AA;

Query Match 100.0%; Score 67; DB 21; Length 162;
 Best Local Similarity 100.0%; Pred. No. 0.00051;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFLERFKSLQKMI 14
 Db 135 eflerfksllqkmi 148

RESULT 3

B18627
 ID B18627 standard; Protein; 519 AA.

XX B18627;

XX 22-JAN-2001 (first entry)

XX Amino acid sequence of MBP-human zalphall ligand fusion in pTAP126.

XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.

XX Synthetic.

XX Homo sapiens.

XX WO200053761-A2.

XX 14-SEP-2000.

XX 09-MAR-2000; 2000WO-US06067.

XX 09-MAR-1999; 99US-0264908.

XX 11-MAR-1999; 99US-0265992.

XX 01-JUL-1999; 99US-0142013.

XX (Zymo) ZYMOGENETICS INC.

XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;

XX WPI; 2000-565600/52.

XX N-PSDB; A75599.

XX New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -

XX Example 31; Page 233-235; 256pp; English.

XX The present sequence represents a MBP-human zalphall ligand fusion in
 CC the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
 CC is useful for stimulating the proliferation and development of
 CC haematopoietic cells in vitro and in vivo. Zalphall ligand

CC polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.

XX Sequence 519 AA;

Query Match 100.0%; Score 67; DB 21; Length 519;

Best Local Similarity 100.0%; Pred. No. 0.0017; Indels 0; Gaps 0;
 Matches 14; Conservative 0; Mismatches 0;

QY 1 EFLERFKSLLOKMI 14
 |||||
 Db 492 eflerfkllqkmi 505

RESULT 4

BI8624 BI8624 standard; Protein; 146 AA.

XX AC BI8624;

XX DT 22-JAN-2001 (first entry)

XX DE A mouse zalphall ligand polypeptide.

XX KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.

XX OS Mus musculus.

XX PN WO200053761-A2.

XX PD 14-SEP-2000.

XX PF 09-MAR-2000; 2000WO-US06067.

XX PR 09-MAR-1999; 99US-0264908.

XX PR 11-MAR-1999; 99US-0265992.

XX PR 01-JUL-1999; 99US-0142013.

XX PA (Zymo) ZYMOGENETICS INC.

XX PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;

XX WPI: 2000-565600/52.
 DR N-PSDB; A75580.

PT New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -

XX PS Disclosure; Page 222-223; 256pp; English.

CC The present sequence represents a mouse zalphall ligand polypeptide,
 CC which is a cytokine. The zalphall ligand is useful for stimulating the
 CC proliferation and development of haematopoietic cells in vitro and in
 CC vivo. Zalphall ligand polynucleotides can be used as primers or probes
 CC for cloning the zalphall gene. The zalphall ligand is useful for
 CC treating tumourigenesis. A zalphall ligand-saporin fusion toxin may be
 CC used for treating leukaemias and lymphomas. Antagonists against zalphall
 CC ligand are useful as research reagents for characterizing ligand-receptor

CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to
 CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.

XX Sequence 146 AA;

Query Match 80.68; Score 54; DB 21; Length 146;

Best Local Similarity 85.7%; Pred. No. 0.069; Indels 0; Gaps 0;
 Matches 12; Conservative 0; Mismatches 2;

QY 1 EFLERFKSLLOKMI 14
 |||||
 Db 128 eflerfkllqkmi 141

RESULT 5

BI8628 BI8628 standard; Protein; 510 AA.

XX AC BI8628;

XX DT 22-JAN-2001 (first entry)

XX DE Amino acid sequence of MBP-mouse zalphall ligand fusion in pTAP134.

XX KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
 KW tumourigenesis; leukaemia; hematopoiesis; B cell tumour.

XX OS Synthetic.

XX OS Mus musculus.

XX PN WO200053761-A2.

XX PD 14-SEP-2000.

XX PF 09-MAR-2000; 2000WO-US06067.

XX PR 09-MAR-1999; 99US-0264908.

XX PR 11-MAR-1999; 99US-0265992.

XX PR 01-JUL-1999; 99US-0142013.

XX PA (Zymo) ZYMOGENETICS INC.

XX PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
 PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;

XX WPI: 2000-565600/52.
 DR N-PSDB; A75602.

PT New human cytokine, designated zalphall ligand, useful for stimulating
 PT the proliferation and/or development of haematopoietic cells in vitro
 PT and in vivo, and for treating tumourigenesis -

XX PS Example 31; Page 239-240; 256pp; English.

CC The present sequence represents a MBP-mouse zalphall ligand fusion in
 CC the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
 CC is useful for stimulating the proliferation and development of
 CC haematopoietic cells in vitro and in vivo. Zalphall ligand
 CC polynucleotides can be used as primers or probes for cloning the
 CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
 CC A zalphall ligand-saporin fusion toxin may be used for treating
 CC leukaemias and lymphomas. Antagonists against zalphall ligand are
 CC useful as research reagents for characterizing ligand-receptor
 CC interaction. Antagonists are also useful for inhibiting expansion,
 CC proliferation, activation and differentiation of cells involved in
 CC regulating hematopoiesis. The zalphall ligand may also be used to

CC stimulate an immune response against B cell tumour, a virus, a parasite
 CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
 CC agonists and antibodies are also useful for the detection, diagnosis,
 CC prevention, and treatment of diseases associated with a zalphall ligand
 CC genetic defect.
 XX
 XX

SQ Sequence 510 AA;

Query Match 80.6%; Score 54; DB 21; Length 510;
 Best Local Similarity 85.7%; Pred. No. 0.25;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 EFLERFKSLQKMI 14

Db 492 eflerlkwllqkmi 505

RESULT 6

ID W27117 standard; Protein; 133 AA.

XX AC W27117;

XX 20-NOV-1997 (first entry)
 XX Feline interleukin-4.

XX Vaccine adjuvant; immunomodulator; tumour; monoclonal antibody;
 XX cat.
 XX Felis domesticus.

XX EP759468-A1.

XX 26-FEB-1997.

XX 10-AUG-1995; 95EP-0401879.
 XX 10-AUG-1995; 95EP-0401879.

XX (VIRB-) LAB VIRBAC.
 XX Horzinek MC, Schijns VECJ;

XX WPI; 1997-147517/14.
 XX N-ESDB; T85127.

XX Feline interleukin-4 protein and DNA sequences - useful as vaccine
 PT adjuvant and immuno-modulator for treating infectious diseases and
 PT tumours in feline(s)

XX Claim 2; Fig 1; 20pp; English.

XX The present sequence represents feline interleukin-4 (fIL-4). The fIL-4
 CC is useful as a vaccine adjuvant and as a therapeutic immunomodulator
 CC for treating infectious diseases and tumours in felines. A mutant
 CC (preferably with the replacement of Tyr-106 by Asp, or with a single
 CC point mutation at one of the last 20 amino acid positions) is useful
 CC as an antagonist that binds to cellular fIL-4 receptors but fails to
 CC mediate signal transduction.

SQ Sequence 133 AA;

Query Match 62.7%; Score 42; DB 18; Length 133;
 Best Local Similarity 58.3%; Pred. No. 6.4;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 EFLERFKSLQK 12

Db 117 dflerlkaimqk 128

RESULT 7

G14687

XX G14687 standard; Protein; 277 AA.

XX AC G14687;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 14644.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0125548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 23-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 30-APR-1999; 99US-0132407.

XX 04-MAY-1999; 99US-0132484.

XX 05-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 07-MAY-1999; 99US-0132863.

XX 11-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

XX 14-MAY-1999; 99US-0134370.

XX 18-MAY-1999; 99US-0134768.

XX 19-MAY-1999; 99US-0134941.

XX 20-MAY-1999; 99US-0135124.

XX 21-MAY-1999; 99US-0135353.

XX 24-MAY-1999; 99US-0135629.

XX 25-MAY-1999; 99US-0136021.

XX 27-MAY-1999; 99US-0136392.

XX 28-MAY-1999; 99US-0136782.

XX 01-JUN-1999; 99US-0137222.

XX 03-JUN-1999; 99US-0137528.

XX 04-JUN-1999; 99US-0137502.

XX 07-JUN-1999; 99US-0137724.

XX 08-JUN-1999; 99US-0138094.

XX 10-JUN-1999; 99US-0138540.

XX 10-JUN-1999; 99US-0138847.

XX 14-JUN-1999; 99US-0139119.

XX 16-JUN-1999; 99US-0139452.

XX 16-JUN-1999; 99US-0139453.

XX 17-JUN-1999; 99US-0139492.

XX 18-JUN-1999; 99US-0139454.

XX 18-JUN-1999; 99US-0139455.

XX 18-JUN-1999; 99US-0139456.

XX 18-JUN-1999; 99US-0139457.

PR	20-AUG-1999;	9905-0149729;
PR	20-AUG-1999;	9905-0149929;
PR	23-AUG-1999;	9905-0149902;
PR	23-AUG-1999;	9905-0149930;
PR	25-AUG-1999;	9905-0150566;
PR	26-AUG-1999;	9905-0150684;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151066;
PR	30-AUG-1999;	9905-0151080;
PR	30-AUG-1999;	9905-0151303;
PR	31-AUG-1999;	9905-0151438;
PR	01-SEP-1999;	9905-0151930;
PR	07-SEP-1999;	9905-0152363;
PR	10-SEP-1999;	9905-0153070;
PR	13-SEP-1999;	9905-0153758;
PR	15-SEP-1999;	9905-0154018;
PR	16-SEP-1999;	9905-0154039;
PR	20-SEP-1999;	9905-0154779;
PR	22-SEP-1999;	9905-0155139;
PR	23-SEP-1999;	9905-0155486;
PR	24-SEP-1999;	9905-0155659;
PR	28-SEP-1999;	9905-0156458;
PR	29-SEP-1999;	9905-0156596;
PR	04-OCT-1999;	9905-0157117;
PR	05-OCT-1999;	9905-0157753;
PR	06-OCT-1999;	9905-0157865;
PR	07-OCT-1999;	9905-0158029;
PR	08-OCT-1999;	9905-0158232;
PR	12-OCT-1999;	9905-0158369;
PR	13-OCT-1999;	9905-0159293;
PR	13-OCT-1999;	9905-0159294;
PR	13-OCT-1999;	9905-0159295;
PR	14-OCT-1999;	9905-0159329;
PR	14-OCT-1999;	9905-0159330;
PR	14-OCT-1999;	9905-0159331;
PR	14-OCT-1999;	9905-0159637;
PR	14-OCT-1999;	9905-0159638;
PR	18-OCT-1999;	9905-0159584;
PR	21-OCT-1999;	9905-0160741;
PR	21-OCT-1999;	9905-0160767;
PR	21-OCT-1999;	9905-0160768;
PR	21-OCT-1999;	9905-0160770;
PR	21-OCT-1999;	9905-0160814;
PR	21-OCT-1999;	9905-0160815;
PR	22-OCT-1999;	9905-0160815;
PR	22-OCT-1999;	9905-0160980;
PR	22-OCT-1999;	9905-0160981;
PR	22-OCT-1999;	9905-0160989;
PR	25-OCT-1999;	9905-0161404;
PR	25-OCT-1999;	9905-0161405;
PR	25-OCT-1999;	9905-0161406;
PR	26-OCT-1999;	9905-0161359;
PR	26-OCT-1999;	9905-0161360;
PR	26-OCT-1999;	9905-0161361;
PR	28-OCT-1999;	9905-0161920;
PR	28-OCT-1999;	9905-0161992;
PR	28-OCT-1999;	9905-0161993;
PR	29-OCT-1999;	9905-0162142;
PR	29-OCT-1999;	9905-0162142;

Best local similarity	100.0%	Freq. NO. 29,
Matches	8; Conservative	0; Mismatches
		0; Indels
		0; Gaps

XX

PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 59.7%; Score 40; DB 21; Length 304;

Best Local Similarity 100.08; Pred. No. 32;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFLERFKS 8

Db 160 eflerfks 167
|||||||

RESULT 9

GI4685
ID GI4685 standard; Protein; 305 AA.

XX AC GI4685;

XX AC

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 14642.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 15-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.

```
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145921.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.

PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 59.7%; Score 40; DB 21; Length 305;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFLERFKS 8
Db 161 eflerfks 168

RESULT 10
R92428
ID R92428 standard; peptide; 23 AA.
XX
AC R92428;
XX
DT 18-SEP-1996 (first entry)
XX
DE Lytic peptide used in ubiquitin-lytic peptide fusion protein.
XX
KW Ubiquitin; fusion protein; lysis; infection; neoplasia; wound healing;
XX
XX stability; reduced toxicity.
OS Synthetic.
XX
XX WO9603519-A1.
XX
PD 08-FEB-1996.
XX
PF 24-JUL-1995; 95WO-US09339.
```

XX 22-JUL-1994; 94US-0279472.
 XX (DEME-) DEMETER BIOTECHNOLOGIES LTD.
 PA (USDA) US SEC OF AGRIC.
 XX Belknap W, Garbarino J, Jaynes J;
 XX WPI; 1996-117061/12.
 XX New fusion protein of ubiquitin and a lytic peptide - for treating
 PT infections and neoplasia, heating wounds, etc. also related nucleic
 PT acid, vectors, and transformed cells
 XX
 PS Claim 5; Page 25; 112pp; English.
 XX R92372-R92462 are lytic peptides used to create ubiquitin-lytic
 CC peptide fusion proteins in which the ubiquitin polypeptide is linked
 CC at its 3'-terminus to the lytic peptide. The lytic peptides are pref.
 CC selected from either the cecropins, defensins, sarcotoxins, mellitin
 CC and magainins. The fusion proteins (FPs) are useful for treating
 CC protozoal, bacterial, fungal and viral infections and neoplasia (in
 CC plants and animals) in the same way as the FP alone, they also
 CC promote wound healing. FPs produced in bacteria may be cleaved in
 CC vitro by ubiquitin hydrolases to recover the active lytic peptide.
 CC FPs produced in eukaryotic cells are cleaved by endogenous enzymes
 CC to yield lytic peptide. Recombinant DNA encoding the FPs have
 CC greater stability in bacteria than DNA encoding the lytic peptide
 CC only.
 XX Sequence 23 AA;
 SQ

Query Match 58.2%; Score 39; DB 17; Length 23;
 Best Local Similarity 50.0%; Pred. No. 3.4;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLLOKMI 14
 :||:|||| :|||
 Db 4 kflkrfkfkvrkfi 17

RESULT 11
 R89985
 ID R89985 standard; peptide; 23 AA.
 XX R89985;
 AC
 XX 16-SEP-1996 (first entry)
 DT Synthetic lytic peptide #19.
 XX
 DE Lytic peptide; ubiquitin; synthetic analogue; cell membrane; cell lysis;
 KW microbial pathogen; disease-resistant plant; bacterial infection; fungus;
 KW protozoa; virus; neoplasia; fusion protein; hydrolase.
 XX
 OS Synthetic.
 XX
 PN WO9603522-A1.
 XX
 PD 08-FEB-1996.
 XX
 PF 24-JUL-1995; 95WO-US09338.
 XX
 PR 22-JUL-1994; 94US-0279472.
 XX
 PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.
 XX Jaynes J;
 PI WPI; 1996-117064/12.
 XX
 DR Lytic peptide(s), useful for developing disease-resistant plants -

PT can be expressed as fusion protein with ubiquitin for stable prodn.
 PT in bacterial host cells
 XX
 PS Claim 1; Page 78; 111pp; English.
 XX R89967-R90021 and R90726-R90763 represent synthetic analogues of
 CC naturally occurring lytic peptides. Lytic peptides destroy prokaryotic
 CC and other non-host cells by disrupting the cell membrane and promoting
 CC cell lysis. Synthetic lytic peptide analogues have similar or higher
 CC levels of lytic activity for many different types of cells, compared to
 CC naturally occurring forms. The concentration of the synthetic analogue
 CC required to lyse microbial pathogens does not lyse normal mammalian
 CC cells. The lytic peptides can be expressed in plants to allow for the
 CC development of disease-resistant plants. The peptides are useful in
 CC promoting wound healing and combatting bacterial infections in plants.
 CC The lytic peptides can also be used for combatting protozoal, fungal,
 CC viral or bacterial infections or neoplasias in mammals and plants. Lytic
 CC peptide-ubiquitin fusion proteins are suitable for production in
 CC bacterial hosts. Bacteria lack the hydrolase which cleaves the peptide
 CC from ubiquitin, and therefore the active (and cytotoxic) lytic peptide
 CC will not be released in the host cells. The recombinantly produced lytic
 CC peptide can be retrieved from the fusion protein by cleavage in vitro.
 XX Sequence 23 AA;
 SQ

Query Match 58.2%; Score 39; DB 17; Length 23;
 Best Local Similarity 50.0%; Pred. No. 3.4;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLLOKMI 14
 :||:|||| :|||
 Db 4 kflkrfkfkvrkfi 17

RESULT 12
 R92429
 ID R92429 standard; peptide; 27 AA.
 XX R92429;
 AC
 XX 18-SEP-1996 (first entry)
 DT Lytic peptide used in ubiquitin-lytic peptide fusion protein.
 DE
 XX Ubiquitin; fusion protein; lysis; infection; neoplasia; wound healing;
 KW stability; reduced toxicity.
 KW
 XX Synthetic.
 OS
 XX WO9603519-A1.
 PN
 XX 08-FEB-1996.
 PD
 XX 24-JUL-1995; 95WO-US09339.
 PF
 XX 22-JUL-1994; 94US-0279472.
 PR
 XX (DEME-) DEMETER BIOTECHNOLOGIES LTD.
 PA (USDA) US SEC OF AGRIC.
 PA
 XX Belknap W, Garbarino J, Jaynes J;
 PI WPI; 1996-117061/12.
 DR
 XX New fusion protein of ubiquitin and a lytic peptide - for treating
 PT infections and neoplasia, heating wounds, etc. also related nucleic
 PT acid, vectors, and transformed cells
 PT
 XX Claim 5; Page 25; 112pp; English.
 PS
 XX R92372-R92462 are lytic peptides used to create ubiquitin-lytic
 CC peptide fusion proteins in which the ubiquitin polypeptide is linked

CC at its 3'-terminus to the lytic peptide. The lytic peptides are pref.
 CC selected from either the cecropins, defensins, sarcotoxins, melittin
 CC and magainins. The fusion proteins (FPs) are useful for treating
 CC protozoal, bacterial, fungal and viral infections and neoplasia (in
 CC plants and animals) in the same way as the FP alone, they also
 CC promote wound healing. FPs produced in bacteria may be cleaved in
 CC vitro by ubiquitin hydrolases to recover the active lytic peptide.
 CC FPs produced in eukaryotic cells are cleaved by endogenous enzymes
 CC to yield lytic peptide. Recombinant DNA encoding the FPs have
 CC greater stability in bacteria than DNA encoding the lytic peptide
 CC only.

XX SQ Sequence 27 AA;

Query Match 58.2%; Score 39; DB 17; Length 27;
 Best Local Similarity 50.0%; Pred. No. 4;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 EFLERFKSLLOKMI 14
 :||:|||| :|||
 Db 4 kflkrfkfkvrkfi 17

RESULT 13

R92430
 ID R92430 standard; peptide; 27 AA.

XX AC R92430;

XX DT 18-SEP-1996 (first entry)

XX DE Lytic peptide used in ubiquitin-lytic peptide fusion protein.

XX KW Ubiquitin; fusion protein; lysis; infection; neoplasia; wound healing;
 XX KW stability; reduced toxicity.

XX OS Synthetic.

XX PN WO9603519-A1.

XX PD 08-FEB-1996.

XX PF 24-JUL-1995; 95WO-US09339.

XX PR 22-JUL-1994; 94US-0279472.

XX PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.
 XX PA (USDA) US SEC OF AGRIC.

XX PI Belknap W, Garbarino J, Jaynes J;

XX DR WPI; 1996-117061/12.

XX PT New fusion protein of ubiquitin and a lytic peptide - for treating
 XX PT infections and neoplasia, heating wounds, etc. also related nucleic
 XX PT acid, vectors, and transformed cells

XX PS Claim 5; Page 25; 112pp; English.

XX CC R92372-R92462 are lytic peptides used to create ubiquitin-lytic
 CC peptide fusion proteins in which the ubiquitin polypeptide is linked
 CC at its 3'-terminus to the lytic peptide. The lytic peptides are pref.
 CC selected from either the cecropins, defensins, sarcotoxins, melittin
 CC and magainins. The fusion proteins (FPs) are useful for treating
 CC protozoal, bacterial, fungal and viral infections and neoplasia (in
 CC plants and animals) in the same way as the FP alone, they also
 CC promote wound healing. FPs produced in bacteria may be cleaved in
 CC vitro by ubiquitin hydrolases to recover the active lytic peptide.
 CC FPs produced in eukaryotic cells are cleaved by endogenous enzymes
 CC to yield lytic peptide. Recombinant DNA encoding the FPs have
 CC greater stability in bacteria than DNA encoding the lytic peptide
 CC only.

XX SQ Sequence 27 AA;

Query Match 58.2%; Score 39; DB 17; Length 27;
 Best Local Similarity 50.0%; Pred. No. 4;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 EFLERFKSLLOKMI 14
 :||:|||| :|||
 Db 8 kflkrfkfkvrkfi 21

RESULT 14

R89986
 ID R89986 standard; peptide; 27 AA.

XX AC R89986;

XX DT 16-SEP-1996 (first entry)

XX DE Synthetic lytic peptide #20.

XX KW Lytic peptide; ubiquitin; synthetic analogue; cell membrane; cell lysis;
 XX KW microbial pathogen; disease-resistant plant; bacterial infection; fungus;
 XX KW protozoa; virus; neoplasia; fusion protein; hydrolase.

XX OS Synthetic.

XX PN WO9603522-A1.

XX PD 08-FEB-1996.

XX PF 24-JUL-1995; 95WO-US09338.

XX PR 22-JUL-1994; 94US-0279472.

XX PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.

XX PI Jaynes J;

XX DR WPI; 1996-117064/12.

XX PT Lytic peptide(s), useful for developing disease-resistant plants -
 XX PT can be expressed as fusion protein with ubiquitin for stable prodn.
 XX PT in bacterial host cells

XX PS Claim 1; Page 78; 111pp; English.

XX CC R89967-R90021 and R90726-R90763 represent synthetic analogues of
 CC naturally occurring lytic peptides. Lytic peptides destroy prokaryotic
 CC and other non-host cells by disrupting the cell membrane and promoting
 CC cell lysis. Synthetic lytic peptide analogues have similar or higher
 CC levels of lytic activity for many different types of cells, compared to
 CC naturally occurring forms. The concentration of the synthetic analogue
 CC required to lyse microbial pathogens does not lyse normal mammalian
 CC cells. The lytic peptides can be expressed in plants to allow for the
 CC development of disease-resistant plants. The peptides are useful in
 CC promoting wound healing and combatting bacterial infections in plants.
 CC The lytic peptides can also be used for combatting protozoal, fungal,
 CC viral or bacterial infections or neoplasias in mammals and plants. Lytic
 CC peptide-ubiquitin fusion proteins are suitable for production in
 CC bacterial hosts. Bacteria lack the hydrolase which cleaves the peptide
 CC from ubiquitin, and therefore the active (and cytotoxic) lytic peptide
 CC will not be released in the host cells. The recombinantly produced lytic
 CC peptide can be retrieved from the fusion protein by cleavage in vitro.

XX SQ Sequence 27 AA;

Query Match 58.2%; Score 39; DB 17; Length 27;
 Best Local Similarity 50.0%; Pred. No. 4;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Search completed: May 23, 2001, 11:11:39
Job time: 182 sec

OY 1 EFLERPKSLQKMI 14
Db 4 kflkrfkfkvrkfi 17

RESULT 15
R89987
ID R89987 standard; peptide; 27 AA.
XX AC R89987;
XX DT 16-SEP-1996 (first entry)
XX DE Synthetic lytic peptide #21.
XX KW Lytic peptide; ubiquitin; synthetic analogue; cell membrane; cell lysis;
KW microbial pathogen; disease-resistant plant; bacterial infection; fungus;
KW protozoa; virus; neoplasia; fusion protein; hydrolase.
XX OS Synthetic.
XX PN WO9603522-A1.
XX PD 08-FEB-1996.
XX PF 24-JUL-1995; 95WO-US09338.
XX PR 22-JUL-1994; 94US-0279472.
XX PA (DEME-) DEMETER BIOTECHNOLOGIES LTD.
XX PI Jaynes J;
XX DR WPI; 1996-117064/12.
XX PT Lytic peptide(s), useful for developing disease-resistant plants -
PT can be expressed as fusion protein with ubiquitin for stable prodn.
PT in bacterial host cells
XX PS Claim 1; Page 79; 111pp; English.
XX CC R89967-R90021 and R90726-R90763 represent synthetic analogues of
CC naturally occurring lytic peptides. Lytic peptides destroy prokaryotic
CC and other non-host cells by disrupting the cell membrane and promoting
CC cell lysis. Synthetic lytic peptide analogues have similar or higher
CC levels of lytic activity for many different types of cells, compared to
CC naturally occurring forms. The concentration of the synthetic analogue
CC required to lyse microbial pathogens does not lyse normal mammalian
CC cells. The lytic peptides can be expressed in plants to allow for the
CC development of disease-resistant plants. The peptides are useful in
CC promoting wound healing and combatting bacterial infections in plants.
CC The lytic peptides can also be used for combatting protozoal, fungal,
CC viral or bacterial infections or neoplasias in mammals and plants. Lytic
CC peptide-ubiquitin fusion proteins are suitable for production in
CC bacterial hosts. Bacteria lack the hydrolase which cleaves the peptide
CC from ubiquitin, and therefore the active (and cytotoxic) lytic peptide
CC will not be released in the host cells. The recombinantly produced lytic
CC peptide can be retrieved from the fusion protein by cleavage in vitro.
XX SQ Sequence 27 AA;

Query Match 58.2%; Score 39; DB 17; Length 27;
Best Local Similarity 50.0%; Pred. No. 4;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
OY 1 EFLERPKSLQKMI 14
Db 8 kflkrfkfkvrkfi 21

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:09:44 ; Search time 58.85 Seconds
(without alignments)
4.570 Million cell updates/sec

Title: US-09-522-217-2_COPY_135_148

Perfect score: 67

Sequence: 1 EFLERFKSLLOKMI 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	58.2	23	2	US-08-505-486-57
2	39	58.2	23	3	US-08-801-028-57
3	39	58.2	23	3	US-09-340-154-57
4	39	58.2	23	5	PCT-US95-09338-57
5	39	58.2	23	5	PCT-US95-09339-57
6	39	58.2	27	2	US-08-505-486-58
7	39	58.2	27	2	US-08-505-486-59
8	39	58.2	27	3	US-08-801-028-58
9	39	58.2	27	3	US-08-801-028-59
10	39	58.2	27	3	US-09-340-154-58
11	39	58.2	27	3	US-09-340-154-59
12	39	58.2	27	5	PCT-US95-09338-58
13	39	58.2	27	5	PCT-US95-09338-59
14	39	58.2	27	5	PCT-US95-09339-58
15	39	58.2	27	5	PCT-US95-09339-59
16	36	53.7	22	3	US-08-940-095-17
17	36	53.7	22	3	US-08-940-095-23
18	36	53.7	22	3	US-08-940-095-136
19	36	53.7	22	3	US-08-940-093-17
20	36	53.7	22	3	US-08-940-093-23
21	36	53.7	22	3	US-08-940-093-136
22	36	53.7	22	3	US-08-940-096-17
23	36	53.7	22	3	US-08-940-096-23
24	36	53.7	22	3	US-08-940-096-136
25	36	53.7	22	3	US-08-884-681-5
26	35	52.2	19	6	5304631-10
27	35	52.2	170	3	US-09-130-663-2
28					Sequence 2, Appl1

28	35	52.2	170	4	US-09-432-335-2	Sequence 2, Appl1
29	35	52.2	292	2	US-08-879-260-2	Sequence 2, Appl1
30	35	52.2	512	2	US-08-194-981E-5	Sequence 5, Appl1
31	35	52.2	1049	3	US-08-772-270A-11	Sequence 11, Appl1
32	35	52.2	1244	5	PCT-US93-10500-2	Sequence 2, Appl1
33	34.5	51.5	416	3	US-08-554-385-21	Sequence 21, Appl1
34	34	50.7	22	3	US-08-940-095-169	Sequence 169, App
35	34	50.7	22	3	US-08-940-095-171	Sequence 171, App
36	34	50.7	22	3	US-08-940-095-182	Sequence 182, App
37	34	50.7	22	3	US-08-940-095-183	Sequence 183, App
38	34	50.7	22	3	US-08-940-093-169	Sequence 169, App
39	34	50.7	22	3	US-08-940-093-171	Sequence 171, App
40	34	50.7	22	3	US-08-940-093-182	Sequence 182, App
41	34	50.7	22	3	US-08-940-093-183	Sequence 183, App
42	34	50.7	22	3	US-08-940-096-169	Sequence 169, App
43	34	50.7	22	3	US-08-940-096-171	Sequence 171, App
44	34	50.7	22	3	US-08-940-096-182	Sequence 182, App
45	34	50.7	22	3	US-08-940-096-183	Sequence 183, App

ALIGNMENTS

RESULT 1

US-08-505-486-57

; Sequence 57, Application US/08505486

; Patent No. 5955573

; GENERAL INFORMATION:

; APPLICANT: Jesse M. Jaynes

; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE

; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND METHODS OF MAKING AND USING SAME

; NUMBER OF SEQUENCES: 98

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ

; STREET: 555 Thirteenth Street N.W.

; CITY: Washington

; STATE: D. C.

; COUNTRY: U.S.A.

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE

; COMPUTER: IBM COMPATIBLE

; OPERATING SYSTEM: DOS

; SOFTWARE: WordPerfect 5.1+

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/505,486

; FILING DATE: 21-JUL-1995

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: U.S. 08/279,472

; FILING DATE: 22-JUL-1994

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: WALKER, BARBARA W.

; REGISTRATION NUMBER: 35,400

; REFERENCE/DOCKET NUMBER: 2093-117A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)783-6040

; TELEFAX: (202)783-6031

; INFORMATION FOR SEQ ID NO: 57:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 23

; TYPE: AMINO ACID

; TOPOLOGY: LINEAR

; MOLECULE TYPE:

; DESCRIPTION: PEPTIDE

; HYPOTHETICAL: NO

; FRAGMENT TYPE: COMPLETE PEPTIDE

; ORIGINAL SOURCE: SYNTHETIC

; IMMEDIATE SOURCE: SYNTHETIC

; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED

; US-08-505-486-57

Query Match 58.2%; Score 39; DB 2; Length 23;
Best Local Similarity 50.0%; Pred. No. 4;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLQKMI 14
:||:||||:|
Db 4 KFLKRFKFKVRKFI 17

RESULT 2

US-08-801-028-57
; Sequence 57, Application US/08801028
; Patent No. 6018102
; GENERAL INFORMATION:
; APPLICANT: JOAN GARBARINO
; APPLICANT: JESSE M. JAYNES
; APPLICANT: WILLIAM BELKNAP
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE CONSTRUCTS, PROTEIN PRODUCTION METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STEVEN J. HULTQUIST
; ADDRESS: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
; STREET: 200 PARK DRIVE, SUITE 210
; STREET: P.O. BOX 14329
; CITY: RESEARCH TRIANGLE PARK
; STATE: NORTH CAROLINA
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: APPLE MACINTOSH
; OPERATING SYSTEM: MACINTOSH
; SOFTWARE: M.S. WORD 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,028
; FILING DATE: 19-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/279,472
; FILING DATE: JULY 22, 1994
; APPLICATION NUMBER: 08/225,476
; FILING DATE: 04-20-94
; APPLICATION NUMBER: 08/225,476
; FILING DATE: 04-08-94
; APPLICATION NUMBER: 08/039,620
; FILING DATE: 06-04-93
; APPLICATION NUMBER: 08/148,491
; FILING DATE: 11-08-93
; APPLICATION NUMBER: 08/148,889
; FILING DATE: 11-08-93
; ATTORNEY/AGENT INFORMATION:
; NAME: WASSERMAN, FRAN S.
; REGISTRATION NUMBER: 34,273
; REFERENCE/DOCKET NUMBER: 4013-104
; TELEPHONE: (919)990-9531
; TELEFAX: (919)990-9532
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-801-028-57

Query Match 58.2%; Score 39; DB 3; Length 23;
Best Local Similarity 50.0%; Pred. No. 4;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLQKMI 14
:||:||||:|
Db 4 KFLKRFKFKVRKFI 17

RESULT 3

US-09-340-154-57
; Sequence 57, Application US/09340154
; Patent No. 6084156
; GENERAL INFORMATION:
; APPLICANT: Jesse M. Jaynes
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
; STREET: 555 Thirteenth Street N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: IBM COMPATIBLE
; OPERATING SYSTEM: DOS
; SOFTWARE: Wordperfect 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/340,154
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/505,486
; FILING DATE: 21-JUL-1995
; APPLICATION NUMBER: U.S. 08/279,472
; FILING DATE: 22-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WALKER, BARBARA W.
; REGISTRATION NUMBER: 35,400
; REFERENCE/DOCKET NUMBER: 2093-117A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)783-6040
; TELEFAX: (202)783-6031
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-09-340-154-57

Query Match 58.2%; Score 39; DB 3; Length 23;
Best Local Similarity 50.0%; Pred. No. 4;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLQKMI 14
:||:||||:|
Db 4 KFLKRFKFKVRKFI 17

RESULT 4

PCT-US95-09338-57

; Sequence 57, Application PC/TUS9509338
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
; NUMBER OF SEQUENCES: 98
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09338
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA: 08/279,472
; APPLICATION NUMBER: 08/279,472
; FILING DATE: 22-JUL-1994
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
; PCT-US95-09338-57

Query Match 58.2%; Score 39; DB 5; Length 23;
Best Local Similarity 50.0%; Pred. No. 4;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLLOKMI 14
:||||| :|||
Db 4 KFLKRFKFKVRKFI 17

RESULT 5
PCT-US95-09339-57
; Sequence 57, Application PC/TUS9509339
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
; NUMBER OF SEQUENCES: 98
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09339
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/279,472
; FILING DATE: 22-JUL-1994
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC

; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
; PCT-US95-09339-57

Query Match 58.2%; Score 39; DB 5; Length 23;
Best Local Similarity 50.0%; Pred. No. 4;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLLOKMI 14
:||||| :|||
Db 4 KFLKRFKFKVRKFI 17

RESULT 6
US-08-505-486-58
; Sequence 58, Application US/08505486
; Patent No. 5955573
; GENERAL INFORMATION:
; APPLICANT: Jesse M. Jaynes
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
; STREET: 555 Thirteenth Street N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: IBM COMPATIBLE
; OPERATING SYSTEM: DOS
; SOFTWARE: WordPerfect 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/505,486
; FILING DATE: 21-JUL-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/279,472
; FILING DATE: 22-JUL-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WALKER, BARBARA W.
; REGISTRATION NUMBER: 35,400
; REFERENCE/DOCKET NUMBER: 2093-117A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)783-6040
; TELEFAX: (202)783-6031
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
; US-08-505-486-58

Query Match 58.2%; Score 39; DB 2; Length 27;
Best Local Similarity 50.0%; Pred. No. 4.6;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLLOKMI 14
:||||| :|||
Db 4 KFLKRFKFKVRKFI 17

RESULT 7
US-08-505-486-59
; Sequence 59, Application US/08505486
; Patent No. 5955573
; GENERAL INFORMATION:
; APPLICANT: Jesse M. Jaynes
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
; METHOD OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
; STREET: 555 Thirteenth Street N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: IBM COMPATIBLE
; OPERATING SYSTEM: DOS
; SOFTWARE: WordPerfect 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/505,486
; FILING DATE: 21-JUL-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/279,472
; FILING DATE: 22-JUL-1994
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WALKER, BARBARA W.
; REGISTRATION NUMBER: 35,400
; REFERENCE/DOCKET NUMBER: 2093-117A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)783-6040
; TELEFAX: (202)783-6031
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-505-486-59

Query Match 58.2%; Score 39; DB 2; Length 27;
Best Local Similarity 50.0%; Pred. No. 4.6;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Qy 1 EFLERFKSLLOKMI 14
:||:||||:|
Db 8 KFLKRFKKFVRKEI 21

RESULT 8
US-08-801-028-58
; Sequence 58, Application US/08801028
; Patent No. 6018102
; GENERAL INFORMATION:
; APPLICANT: JOAN GARBARINO
; APPLICANT: JESSE M. JAYNES
; APPLICANT: WILLIAM BELKNAP
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE CONSTRUCTS, PROTEIN PRODUC
; NUMBER OF SEQUENCES: 98

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STEVEN J. HULTQUIST
; ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
; STREET: 200 PARK DRIVE, SUITE 210
; CITY: P.O. BOX 14329
; CITY: RESEARCH TRIANGLE PARK
; STATE: NORTH CAROLINA
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
; COMPUTER: APPLE MACINTOSH
; OPERATING SYSTEM: MACINTOSH
; SOFTWARE: M.S. WORD 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/801,028
; FILING DATE: 19-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/279,472
; FILING DATE: JULY 22, 1994
; APPLICATION NUMBER: 08/225,476
; FILING DATE: 04-20-94
; APPLICATION NUMBER: 08/225,476
; FILING DATE: 04-08-94
; APPLICATION NUMBER: 08/039,620
; FILING DATE: 06-04-93
; APPLICATION NUMBER: 08/148,491
; FILING DATE: 11-08-93
; APPLICATION NUMBER: 08/148,889
; FILING DATE: 11-08-93
; ATTORNEY/AGENT INFORMATION:
; NAME: WASSERMAN, FRAN S.
; REGISTRATION NUMBER: 34,273
; REFERENCE/DOCKET NUMBER: 4013-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)990-9531
; TELEFAX: (919)990-9532
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-801-028-58

Query Match 58.2%; Score 39; DB 3; Length 27;
Best Local Similarity 50.0%; Pred. No. 4.6;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Qy 1 EFLERFKSLLOKMI 14
:||:||||:|
Db 4 KFLKRFKKFVRKEI 17

RESULT 9
US-08-801-028-59
; Sequence 59, Application US/08801028
; Patent No. 6018102
; GENERAL INFORMATION:
; APPLICANT: JOAN GARBARINO
; APPLICANT: JESSE M. JAYNES
; APPLICANT: WILLIAM BELKNAP
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE CONSTRUCTS, PROTEIN PRO
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STEVEN J. HULTQUIST

ADDRESSEE: INTELLECTUAL PROPERTY/TECHNOLOGY LAW
STREET: 200 PARK DRIVE, SUITE 210
STREET: P.O. BOX 14329
CITY: RESEARCH TRIANGLE PARK
STATE: NORTH CAROLINA
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: APPLE MACINTOSH
OPERATING SYSTEM: MACINTOSH
SOFTWARE: M.S. WORD 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,028
FILING DATE: 19-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/279,472
FILING DATE: JULY 22, 1994
APPLICATION NUMBER: 08/225,476
FILING DATE: 04-20-94
APPLICATION NUMBER: 08/225,476
FILING DATE: 04-08-94
APPLICATION NUMBER: 08/039,620
FILING DATE: 06-04-93
APPLICATION NUMBER: 08/148,491
FILING DATE: 11-08-93
APPLICATION NUMBER: 08/148,889
FILING DATE: 11-08-93
ATTORNEY/AGENT INFORMATION:
NAME: WASSERMAN, FRAN S.
REGISTRATION NUMBER: 34,273
REFERENCE/DOCKET NUMBER: 4013-104
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)990-9531
TELEFAX: (919)990-9532
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 27
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: NO
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-08-801-028-59

Query Match 58.2%; Score 39; DB 3; Length 27;
Best Local Similarity 50.0%; Pred. No. 4.6;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERKSLLOKMI 14
:||||| :||
Db 8 KFLRFRKFKVRKFI 21

RESULT 10
US-09-340-154-58
Sequence 58, Application US/09340154
Patent No. 6084156
GENERAL INFORMATION:
APPLICANT: Jesse M. Jaynes
TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
TITLE OF INVENTION: METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
STREET: 555 Thirteenth Street N.W.
CITY: Washington

STATE: D. C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: DOS
SOFTWARE: WordPerfect 5.1+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/340,154
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/505,486
FILING DATE: 21-JUL-1995
APPLICATION NUMBER: U.S. 08/279,472
FILING DATE: 22-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: WALKER, BARBARA W.
REGISTRATION NUMBER: 35,400
REFERENCE/DOCKET NUMBER: 2093-117A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)783-6040
TELEFAX: (202)783-6031
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 27
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: NO
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
US-09-340-154-58

Query Match 58.2%; Score 39; DB 3; Length 27;
Best Local Similarity 50.0%; Pred. No. 4.6;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERKSLLOKMI 14
:||||| :||
Db 4 KFLRFRKFKVRKFI 17

RESULT 11
US-09-340-154-59
Sequence 59, Application US/09340154
Patent No. 6084156
GENERAL INFORMATION:
APPLICANT: Jesse M. Jaynes
TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
TITLE OF INVENTION: METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
STREET: 555 Thirteenth Street N.W.
CITY: Washington
STATE: D. C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.4 MB STORAGE
COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: DOS
SOFTWARE: WordPerfect 5.1+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/340,154
FILING DATE:


```

; CLASSIFICATION:
; PRIOR APPLICATION DATA: 08/505,486
; APPLICATION NUMBER: 08/505,486
; FILING DATE: 21-JUL-1995
; APPLICATION NUMBER: U.S. 08/279,472
; FILING DATE: 22-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WALKER, BARBARA W.
; REGISTRATION NUMBER: 35,400
; REFERENCE/DOCKET NUMBER: 2093-117A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)783-6040
; TELEFAX: (202)783-6031
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; DESCRIPTION: PEPTIDE
; MOLECULE TYPE: NO
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
;
US-09-340-154-59

Query Match 58.2%; Score 39; DB 3; Length 27;
Best Local Similarity 50.0%; Pred. No. 4.6;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLQKMI 14
DB 8 KFLKRFKKFVRKFI 21

RESULT 12
PCT-US95-09338-58
; Sequence 58, Application PC/TUS9509338
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
; NUMBER OF SEQUENCES: 98
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09338
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA: 08/279,472
; APPLICATION NUMBER: 08/279,472
; FILING DATE: 22-JUL-1994
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
;
PCT-US95-09338-59

Query Match 58.2%; Score 39; DB 5; Length 27;
Best Local Similarity 50.0%; Pred. No. 4.6;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLQKMI 14
DB 8 KFLKRFKKFVRKFI 21

RESULT 14
PCT-US95-09339-58
; Sequence 58, Application PC/TUS9509339
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
; NUMBER OF SEQUENCES: 98
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09339
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA: 08/279,472
; APPLICATION NUMBER: 08/279,472
; FILING DATE: 22-JUL-1994
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
;
PCT-US95-09338-58

Query Match 58.2%; Score 39; DB 5; Length 27;
Best Local Similarity 50.0%; Pred. No. 4.6;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLQKMI 14
DB 8 KFLKRFKKFVRKFI 21

RESULT 14
PCT-US95-09339-58
; Sequence 58, Application PC/TUS9509339
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
; NUMBER OF SEQUENCES: 98
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09339
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA: 08/279,472
; APPLICATION NUMBER: 08/279,472
; FILING DATE: 22-JUL-1994
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
;
PCT-US95-09338-59

Query Match 58.2%; Score 39; DB 5; Length 27;
Best Local Similarity 50.0%; Pred. No. 4.6;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLQKMI 14
DB 8 KFLKRFKKFVRKFI 21

RESULT 13
PCT-US95-09338-59
; Sequence 59, Application PC/TUS9509338
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
; TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
; NUMBER OF SEQUENCES: 98
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09338
; FILING DATE: 21-JUL-1994
; PRIOR APPLICATION DATA: 08/279,472
; APPLICATION NUMBER: 08/279,472
; FILING DATE: 22-JUL-1994
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: NO
; FRAGMENT TYPE: COMPLETE PEPTIDE
; ORIGINAL SOURCE: SYNTHETIC
; IMMEDIATE SOURCE: SYNTHETIC
; PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
;
PCT-US95-09338-59
```

INFORMATION FOR SEQ ID NO: 58;
SEQUENCE CHARACTERISTICS:
LENGTH: 27
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: NO
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
PCT-US95-09339-58

Query Match 58.2%; Score 39; DB 5; Length 27;
Best Local Similarity 50.0%; Pred. No. 4.6;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLLOKMI 14
Db 4 KFLKRFKFKVRKFI 17

RESULT 15
PCT-US95-09339-59
Sequence 59, Application PC/TUS9509339
GENERAL INFORMATION:
APPLICANT:
APPLICANT:
TITLE OF INVENTION: UBIQUITIN-LYTIC PEPTIDE FUSION GENE
TITLE OF INVENTION: CONSTRUCTS, PROTEIN PRODUCTS DERIVING THEREFROM, AND
TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
NUMBER OF SEQUENCES: 98
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1+
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09339
FILING DATE: 21-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/279,472
FILING DATE: 22-JUL-1994
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 27
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE:
DESCRIPTION: PEPTIDE
HYPOTHETICAL: NO
FRAGMENT TYPE: COMPLETE PEPTIDE
ORIGINAL SOURCE: SYNTHETIC
IMMEDIATE SOURCE: SYNTHETIC
PUBLICATION INFORMATION: NOT PREVIOUSLY PUBLISHED
PCT-US95-09339-59

Query Match 58.2%; Score 39; DB 5; Length 27;
Best Local Similarity 50.0%; Pred. No. 4.6;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLLOKMI 14
Db 8 KFLKRFKFKVRKFI 21

Search completed: May 23, 2001, 11:09:44
Job time: 67 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:12:54 ; Search time 70.54 Seconds
(without alignments)
13.639 Million cell updates/sec

Title: US-09-522-217-2_COPY_135_148
Perfect score: 67
Sequence: 1 EFLERFKSLLOKMI 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_67:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

-Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	62.7	455	2 T33572	hypothetical prote
2	42	62.7	751	1 WMXRGB	probable core prot
3	41	61.2	205	2 D69804	conserved hypothet
4	41	61.2	542	2 H64319	hypothetical prote
5	41	61.2	570	2 T27407	hypothetical prote
6	41	61.2	2144	2 T21712	hypothetical prote
7	40	59.7	208	2 E72514	hypothetical prote
8	40	59.7	506	2 T28810	hypothetical prote
9	39	58.2	613	2 F81334	hypothetical prote
10	39	58.2	1817	2 H71611	probable secreted
11	38	56.7	303	2 A24862	Na+/K+-exchanging
12	38	56.7	303	2 I47125	ATPase beta-subuni
13	38	56.7	406	2 I46571	Na+, K+-ATPase - p
14	38	56.7	303	2 H83532	N-succinylglutamat
15	38	56.7	640	2 C72351	hypothetical prote
16	38	56.7	672	2 T30374	probable envelope
17	38	56.7	751	2 T30154	hypothetical prote
18	37	55.2	127	2 T14956	hypothetical prote
19	37	55.2	174	2 B71650	hypothetical prote
20	37	55.2	193	2 T18623	hypothetical prote
21	37	55.2	338	2 T36025	conserved hypothet
22	37	55.2	357	2 S73851	hypothetical prote
23	37	55.2	383	2 S76334	hypothetical prote
24	37	55.2	464	2 T21505	hypothetical prote
25	37	55.2	512	2 S21761	aryl hydrocarbon (
26	37	55.2	558	2 C72391	conserved hypothet
27	37	55.2	600	2 T14463	galactonolactone d
28	37	55.2	605	2 S46833	hypothetical prote
29	37	55.2	608	2 E71859	phosphogluconate d

30	37	55.2	608	2 D64657	phosphogluconate d
31	37	55.2	619	2 H81351	hypothetical prote
32	37	55.2	898	2 S52826	hypothetical prote
33	37	55.2	1188	1 NDECKR	type I site-specif
34	37	55.2	1288	2 T09908	hypothetical prote
35	37	55.2	2279	2 T42531	acetyl-CoA carboxy
36	37	55.2	2628	2 S59413	probable membrane
37	37	55.2	3005	2 T08841	polyprotein - dour
38	36.5	54.5	309	2 T50797	hypothetical prote
39	36	53.7	86	2 F64527	hypothetical prote
40	36	53.7	127	2 A33894	chorismate mutase
41	36	53.7	147	2 C64223	hypothetical prote
42	36	53.7	203	2 E71700	hypothetical prote
43	36	53.7	209	2 A72043	ribosomal protein
44	36	53.7	216	2 H84009	hypothetical prote
45	36	53.7	219	2 A64340	hypothetical prote

ALIGNMENTS

RESULT 1
T33572
hypothetical protein Y59C2A.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C:Accession: T33572
R:Ozersky, P.
submitted to the EMBL Data Library, October 1998
A:Description: The sequence of C. elegans cosmid Y59C2A.
A:Reference number: Z21372
A:Accession: T33572
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-455 <OZE>
A:Cross-references: EMBL:AF099003; PIDN:AAC58743.1; GSPDB:GN000020; CESP:Y59C2A.1
A:Experimental source: strain Bristol N2; clone Y59C2A
C:Genetics:
A:Gene: CESP:Y59C2A.1
A:Map position: 2
A:Introns: 33/2; 60/3; 115/2; 152/3; 170/3; 244/3; 276/1; 302/1; 344/3; 373/1; 418/1;
C:Superfamily: carboxypeptidase

Query Match 62.7%; Score 42; DB 2; Length 455;
Best Local Similarity 80.0%; Pred. No. 20;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0; -
QY 1 EFLERFKSL 10
| | | | |
DB 89 EFLQKFKSL 98
RESULT 2
WMXRGB
probable core protein - human rotavirus B
C:Species: human rotavirus B
A:Note: host Homo sapiens (man)
C:date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
C:Accession: A33093; S06927
R:Sato, S.; Volken, R.H.; Eiden, J.J.
Nucleic Acids Res. 17, 10113, 1989
A:title: The complete nucleic acid sequence of gene segment 3 of the IDIR strain of g
A:Reference number: S06927; MUID:90098788
A:Accession: A33093
A:Status: translation not shown
A:Molecule type: genomic RNA
A:Residues: 1-751 <SAT>
A:Cross-references: EMBL:X16949; NID:g61931; PIDN:CAA34823.1; PID:g61932
C:Genetics:
A:Map position: segment 3
C:Superfamily: rotavirus core protein
C:Keywords: core protein

Query Match 62.7%; Score 42; DB 1; Length 751;
 Best Local Similarity 69.2%; Pred. No. 34;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLQKM 13
 ||| | ||| ||
 Db 738 EFLTRIKSLVKM 750

RESULT 3
 D69804
 conserved hypothetical protein yfir - Bacillus subtilis
 C:Species: Bacillus subtilis
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
 C:Accession: D69804
 R:Kunst, F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertel
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033
 A:Accession: D69804
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-205 <KUN>
 A:Cross-references: GB:299108; GB:AL009126; NID:g2633055; PIDN:CAB12666.1; PID:el182827;
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yfir

Query Match 61.2%; Score 41; DB 2; Length 205;
 Best Local Similarity 69.2%; Pred. No. 13;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 FLERFKSLQKMI 14
 | || | || | |
 Db 134 FVERFSLQKGI 146

RESULT 4
 H64319
 hypothetical protein MJ0159 - Methanococcus jannaschii
 C:Species: Methanococcus jannaschii
 C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C:Accession: H64319
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
 Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
 rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
 A>Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
 A:Reference number: A64300; MUID:96337999
 A:Accession: H64319
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-542 <BUL>
 A:Cross-references: GB:U67473; GB:L77117; NID:g2826256; PIDN:AAB98143.1; PID:gl590916;
 C:Genetics:
 A:Map position: REV163055-161427
 A:Start codon: GTC

Query Match 61.2%; Score 41; DB 2; Length 542;
 Best Local Similarity 63.6%; Pred. No. 35;
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFLERFKSLQ 11
 ||||| : : : : :
 Db 233 EFLERFETILE 243

RESULT 5
 T27407
 hypothetical protein Y75B8A.25 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
 C:Accession: T27407
 R:Barlow, K.
 submitted to the EMBL Data Library, November 1998
 A:Reference number: Z20361
 A:Accession: T27407
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-570 <WIL>
 A:Cross-references: EMBL:AL033514; PIDN:CAA22109.1; CESP:Y75B8A.25
 A:Experimental source: clone Y75B8A
 C:Genetics:
 A:Gene: CESP:Y75B8A.25
 A:Introns: 38/2; 128/2; 327/3; 389/3; 475/1; 518/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein Y75B8A.25

Query Match 61.2%; Score 41; DB 2; Length 570;
 Best Local Similarity 75.0%; Pred. No. 37;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EFLERFKSLQK 12
 | | | | | : : : : :
 Db 343 EILEFFKLLQK 354

RESULT 6
 T21712
 hypothetical protein F33H2.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C:Accession: T21712
 R:Cottage, A.
 submitted to the EMBL Data Library, November 1996
 A:Reference number: Z19463
 A:Accession: T21712
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-2144 <WIL>
 A:Cross-references: EMBL:Z81526; PIDN:CAB04263.1; GSPDB:GN00019; CESP:F33H2.5
 A:Experimental source: clone F33H2
 C:Genetics:
 A:Gene: CESP:F33H2.5
 A:Map position: 1
 A:Introns: 14/2; 227/2; 337/3; 544/2; 626/3; 561/1; 776/1; 964/3; 1153/2; 1720/2; 187
 C:Superfamily: DNA-directed DNA polymerase II

Query Match 61.2%; Score 41; DB 2; Length 2144;
 Best Local Similarity 61.5%; Pred. No. 1.4e+02;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 FLERFKSLQKMI 14
 : || | | | : : : : :
 Db 1112 YLERFGSCIQKII 1124

RESULT 7

E72514
 hypothetical protein APE2090 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: E72514
 R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum A:Reference number: A72450; MUID:99310339
 A:Accession: E72514
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-208 <RAW>
 A:Cross-references: DDBJ:AP000063; NID:g5105654; PIDN:BAA81101.1; PID:g5105789
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE2090
 C:Superfamily: dTMP kinase

Query Match 59.7%; Score 40; DB 2; Length 208;
 Best Local Similarity 50.0%; Pred. No. 20;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 EFLERFKSLLOKMI 14
 ||||| ||: ||:
 Db 151 EFLERKSMYEVL 164

RESULT 8
 T28810
 hypothetical protein C54D1.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T28810
 R:Minx, M.
 submitted to the EMBL Data Library, January 1996
 A:Description: The sequence of C. elegans cosmid C54D1.
 A:Reference number: Z20527
 A:Accession: T28810
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-506 <MIN>
 A:Cross-references: EMBL:U46673; PIDN:AAC48151.1; GSPDB:GN000028; CESP:C54D1.4
 A:Experimental source: strain Bristol N2; clone C54D1
 C:Genetics:
 A:Gene: CESP:C54D1.4
 A:Map position: X
 A:Introns: 78/2; 121/2; 180/3; 296/2; 332/3; 384/3; 421/3; 464/3
 C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology

Query Match 59.7%; Score 40; DB 2; Length 506;
 Best Local Similarity 53.8%; Pred. No. 48;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 FLERFKSLLOKMI 14
 : ||: |||||
 Db 6 YREFKNILQKLI 18

RESULT 9
 F81334
 hypothetical protein Cj1268c [imported] - Campylobacter jejuni (strain NCTC 11168)
 C:Species: Campylobacter jejuni
 C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Mar-2000
 C:Accession: F81334
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrer Nature 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp A:Reference number: A81250; MUID:20150912

A:Accession: F81334
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-613 <PAR>
 A:Cross-references: GB:AL1139077; GB:AL111168; NID:g6968444; PIDN:CAB73522.1; PID:g696 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 A:Gene: Cj1268c

Query Match 58.2%; Score 39; DB 2; Length 613;
 Best Local Similarity 50.0%; Pred. No. 86;
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 EFLERFKSLLOKMI 14
 || || || ||:
 Db 103 EFYEFEKLEQFL 116

RESULT 10
 H71611
 probable secreted protein PFB0565w - malaria parasite (Plasmodium falciparum)
 C:Species: Plasmodium falciparum
 C>Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
 C:Accession: H71611
 R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E. ; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H. Science 282, 1126-1132, 1998
 A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum. A:Reference number: A71600; MUID:99021743
 A:Accession: H71611
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1817 <GAR>
 A:Cross-references: GB:AE001403; GB:AE001362; NID:g3845216; PIDN:AAC71902.1; PID:g384 A:Experimental source: clone 3D7
 C:Genetics:
 A:Gene: PFB0565w

Query Match 58.2%; Score 39; DB 2; Length 1817;
 Best Local Similarity 57.1%; Pred. No. 2.5e+02;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLLOKMI 14
 ||||: |||| |:
 Db 1793 EFLRFRKIYLDIIL 1806

RESULT 11
 A24862
 Na+/K+-exchanging ATPase (EC 3.6.1.37) beta chain - pig
 N:Alternate names: sodium pump beta chain; sodium/potassium-dependent ATPase beta cha C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 22-Jun-1999
 C:Accession: A24862; S00037; S00501; S29761
 R:Ovchinnikov, Y.A.; Modyanov, N.N.; Broude, N.E.; Petrukhin, K.E.; Grishin, A.V.; Ar FEBS Lett. 201, 237-245, 1986
 A:Title: Pig kidney Na+/K+-ATPase. Primary structure and spatial organization.
 A:Reference number: A91361; MUID:86220813
 A:Accession: A24862
 A:Molecule type: mRNA
 A:Residues: 1-303 <OVCC>
 A:Cross-references: EMBL:X03937; NID:q1899; PIDN:CAA27575.1; PID:q1900
 A:Note: the authors translated the codon TCC for residue 151 as Phe
 A:Note: part of this sequence, including the amino and carboxyl end of the mature pro R:Ovchinnikov, Y.A.; Arzamakov, N.M.; Arystarkhova, E.A.; Gevondyan, N.M.; Aldanova, FEBS Lett. 217, 269-274, 1987
 A:Title: Detailed structural analysis of exposed domains of membrane-bound Na+,K+-ATP A:Reference number: S00011; MUID:87247231
 A:Contents: annotation; membrane topology
 C:Superfamily: Na+/K+-transporting ATPase beta chain
 C:Keywords: glycoprotein; heterodimer; hydrolase; ion transport; potassium transport;

RESULT 14

H83532
N-succinylglutamate 5-semialdehyde dehydrogenase PA0895 [imported] - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: H83532
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: AB2950; MUID:20437337
A:Accession: H83532
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-406 <STO>
A:Cross-references: GB:AE004524; GB:AE004091; MID:g9946795; PIDN:AAG04284.1; GSPDB:GN000001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: aruC; PA0895
C:Superfamily: ornithine--oxo-acid aminotransferase

Query Match	56.7%	Score 38;	DB 2;	Length 406;
Best Local Similarity	80.0%;	Pred. No. 83;		
Matches	8;	Conservative	1;	Mismatches
			1;	Indels
				0;
				Gaps
				0;
Oy	4	ERFKSLQKM 13		
Db	316	ERFKSLQKI 325		

RESULT 15

C72351
hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: C72351
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, J.R.; Nelson, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing of *Thermotoga maritima*
A:Reference number: A72200; MUID:99287316
A:Accession: C72351
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-640 <ARN>
A:Cross-references: GB:AE001738; GB:AE000512; MID:g4981158; PIDN:AAD35728.1; PID:g4981158
C:Genetics:
A:Gene: TM0644

Query Match	56.7%	Score 38;	DB 2;	Length 640;
Best Local Similarity	63.6%;	Pred. No. 1.3e+02;		
Matches	7;	Conservative	3;	Mismatches
			1;	Indels
				0;
				Gaps
				0;
Oy	3	LERFKSLQKM 13		
		:: :		
Db	230	IDRYASLLQKM 240		

Search completed: May 23, 2001, 11:12:56
Job time: 259 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:20:27 ; Search time 40.06 Seconds
(without alignments)
11.971 Million cell updates/sec

Title: US-09-522-217-2_COPY_135_148
Perfect score: 67
Sequence: 1 EFLERFKSLQKMI 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	62.7	133	IL4_FELCA	P55030 felis silve
2	42	62.7	751	VPA_ROTGI	P15155 rotavirus (
3	41	61.2	542	Y159_METJA	Q57623 methanococ
4	40	59.7	208	KTHY_AERPE	Q99448 aeropyrum p
5	39	58.2	132	IL4_CANFA	O77762 canis famil
6	38	56.7	303	ATNB_PIG	P05027 sus scrofa
7	37	55.2	136	SCPB_APLSP	P09892 aplysia sp.
8	37	55.2	174	Y875_RICPR	Q92692 rickettsia
9	37	55.2	193	YP2D_CABEL	Q09211 caenorhabdi
10	37	55.2	357	Y21C_MYCPN	P75470 mycoplasma
11	37	55.2	512	CP11_MACEA	P33616 macaca fasc
12	37	55.2	605	APM2_YEAST	P38700 saccharomyc
13	37	55.2	608	EDD_HELPJ	Q924b3 helicobacte
14	37	55.2	608	EDD_HELPY	P56111 helicobacte
15	37	55.2	898	YMW6_YEAST	Q04748 saccharomyc
16	37	55.2	1188	TRK2_COLI	P08956 escherichia
17	37	55.2	1576	RPOC_AQUPY	Q9x5y2 aquifex pyr
18	36	53.7	127	CHMU_BACSU	P19080 bacillus su
19	36	53.7	147	Y211_MYCGE	P47453 mycoplasma
20	36	53.7	209	R54_CHLPN	Q927h2 chlamydia p
21	36	53.7	219	ENGB_METJA	Q57768 methanococ
22	36	53.7	219	Y320_METJA	Q57768 methanococ
23	36	53.7	321	AVRB_PSESG	P13835 pseudomonas
24	36	53.7	323	FRA2_CHICK	P18625 gallus gall
25	36	53.7	349	FPPS_KLULA	P49349 kluyveromyc
26	36	53.7	385	PHEA_BUCAL	P57472 bucinera ap
27	36	53.7	414	Y701_METJA	Q58112 methanococ
28	36	53.7	872	ATCL_MYCPN	P78036 mycoplasma
29	36	53.7	885	ASE1_YEAST	P50275 saccharomyc
30	36	53.7	901	OCRL_HUMAN	Q01968 homo sapien
31	36	53.7	1026	MY1B_DROME	Q23979 drosophila
32	36	53.7	1164	KEL1_YEAST	P38853 saccharomyc
33	35	52.2	66	NXL2_BUNFL	P15815 bungarus fl

RESULT 1
ID IL4_FELCA STANDARD; PRT; 133 AA.
AC P55030; P79170; O62774;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-4 PRECURSOR (IL-4) (B-CELL STIMULATORY FACTOR 1) (BSF-1)
DE (LYMPHOCYTE STIMULATORY FACTOR 1).
GN IL4.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RX [1]
RC SEQUENCE FROM N.A.
RP TISSUE=Peripheal blood;
RA Schijns V.E.C.J., Wierda C.M.H., van Dam E.J.M., Vahlenkamp T.W.,
RA Horzinek M.C.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Lerner D.L., Elder J.H.;
RN Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 25-126 FROM N.A.
RA Harley R., Helps C.R., Grufflydd-Jones T.J., Day M.J., Harbour D.A.;
RN Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: IL-4 PARTICIPATES IN AT LEAST SEVERAL B-CELL ACTIVATION
CC PROCESSES AS WELL AS OF OTHER CELL TYPES. IT IS A COSTIMULATOR OF
CC DNA-SYNTHESIS. IT INDUCES THE EXPRESSION OF CLASS II MHC MOLECULES
CC ON RESTING B-CELLS. IT ENHANCES BOTH SECRETION AND CELL SURFACE
CC EXPRESSION OF ICE AND IGGL. IT ALSO REGULATES THE EXPRESSION OF
CC THE LOW AFFINITY FC RECEPTOR FOR IGE (CD23) ON BOTH LYMPHOCYTES
CC AND MONOCYTES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE IL-4 / IL-13 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X87408; CAA60856.1; -;
DR EMBL; U93634; AAB42052.1; -;
DR EMBL; AF054602; AAC15975.1; -;
DR HSSP; P05112; 1CYL.
DR InterPro; IPR001325; -;
DR InterPro; IPR002354; -;
DR Pfam; PF00727; IL4; 1
DR PRINTS; PR00431; INTERLEUKIN4.
DR PROSITE; PS00838; INTERLEUKIN_4_13; 1.
KW cytokine; Glycoprotein; B-cell activation; Growth factor; Signal.

34 35 52.2 185 1 VATE_THEAC O9hm68 thermoplasm
35 35 52.2 196 1 YQXD_BACSU P17868 bacillus su
36 35 52.2 278 1 HBD_DEIRA Q9rvgl deinococcus
37 35 52.2 292 1 TTC1_HUMAN Q99614 homo sapien
38 35 52.2 369 1 DNAB_NITEU O06431 nitrosomona
39 35 52.2 376 1 UL53_HCMVA P16794 human cytom
40 35 52.2 392 1 HEM1_METJA Q60172 methanococ
41 35 52.2 401 1 Y148_METSM P22344 methanobrev
42 35 52.2 440 1 BRA2_BRAFL P80492 branchiosto
43 482 1 NUSA_BORBU O51740 borrelia bu
44 35 52.2 488 1 SYFA_AERPE O9y9i6 aeropyrum p
45 35 52.2 512 1 CP11_HUMAN P04798 homo sapien

ALIGNMENTS


```
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 133 INTERLEUKIN-4.
FT DISULFID 48 85 BY SIMILARITY.
FT DISULFID 70 113 BY SIMILARITY.
FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 45 45 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 2 2 G -> D (IN REF. 2).
FT CONFLICT 5 5 Y -> S (IN REF. 2).
SQ SEQUENCE 133 AA; 15148 MW; 6D94C71F61EFC75 CRC64;

Query Match 62.7%; Score 42; DB 1; Length 133;
Best Local Similarity 58.3%; Pred. No. 2.7;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EFLERFKSLLOK 12
Db 117 DFLERIKAIMOK 128

RESULT 2
VP4_ROTGI STANDARD; PRT; 751 AA.
ID VP4_ROTGI
AC P15155;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE OUTER CAPSID PROTEIN VP4 (HEMAGGLUTININ) (OUTER LAYER PROTEIN VP4).
GN S3.
OS Rotavirus (group B / strain IDIR) (IDIR agent).
OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX NCBI_TaxID=28877;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90098788; PubMed=2557579;
RA Sato S., Yoken R.H., Eiden J.J.;
RT "The complete nucleic acid sequence of gene segment 3 of the IDIR
strain of group B rotavirus."
RL Nucleic Acids Res. 17:10113-10113(1989).
CC -!- SUBCELLULAR LOCATION: OUTER CAPSID.
CC -!- SIMILARITY: BELONGS TO THE ROTAVIRUS VP4 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X16949; CAA34823.1; -
DR PIR; A33093; WMXRGB.
DR InterPro; IPR000416; -
DR Pfam; PF00426; VP4; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 211 211 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 528 528 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 678 678 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 699 699 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 751 AA; 85557 MW; 756A33227C065709 CRC64;

Query Match 62.7%; Score 42; DB 1; Length 751;
Best Local Similarity 69.2%; Pred. No. 16;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 EFLERFKSLLOK 13
Db 738 EFLTRIKSLVLM 750

RESULT 3
Y159_METJA STANDARD; PRT; 542 AA.
ID Y159_METJA
AC Q57623;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL PROTEIN MJ0159.
GN MJ0159.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii."
RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67473; AAB98143.1; -
DR TIGR; MJ0159; -
DR InterPro; IPR002846; -
DR Pfam; PF01995; DUF128; 2.
KW Hypothetical protein.
SQ SEQUENCE 542 AA; 61195 MW; 5C7FC0734338AE0F CRC64;

Query Match 61.2%; Score 41; DB 1; Length 542;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFLERFKSLLO 11
Db 233 EFLERPETILE 243

RESULT 4
KTHY_AERPE STANDARD; PRT; 208 AA.
ID KTHY_AERPE
AC Q9YA48;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROBABLE THYMIDYLATE KINASE (EC 2.7.4.9) (DTMP KINASE).
GN TMK OR APE2090.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC Aeropyrum.
OX NCBI_TaxID=56636;
```

RN SEQUENCE FROM N.A.
 RC STRAIN-K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 crenarchaeon, Aeropyrum pernix K1.";
 RL DNA Res. 6:83-101(1999).
 CC -!- CATALYTIC ACTIVITY: ATP + THYMIDINE 5'-MONOPHOSPHATE = ADP +
 THYMIDINE 5'-DIPHOSPHATE.
 CC -!- PATHWAY: SYNTHESIS OF DTPP FROM DTPP.
 CC -!- SIMILARITY: BELONGS TO THE THYMIDYLATE KINASE FAMILY.
 CC -!- CAUTION: COULD BE INACTIVE DUE TO A DEFECTIVE ATP-BINDING SITE.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AP000063; BAA81101.1; -
 CC InterPro: IPR000062; -
 CC DR PROSITE; PS01331; THYMIDYLATE KINASE; 1.
 KW Transferase; Kinase; Nucleotide biosynthesis; ATP-binding.
 FT SITE 8 15 DEFECTIVE ATP-BINDING (POTENTIAL).
 SQ SEQUENCE 208 AA; 22945 MW; ED155CE4ECA3382F CRC64;
 Query Match 59.7%; Score 40; DB 1; Length 208;
 Best Local Similarity 50.0%; Pred. No. 9.1;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EFLERFKSLQKMI 14
 Db 151 EFLERKSMYEEVL 164
 RESULT 5
 ID IL4_CANFA STANDARD; PRT; 132 AA.
 AC Q77762;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE INTERLEUKIN-4 PRECURSOR (IL-4) (B-CELL STIMULATORY FACTOR 1) (BSP-1)
 DE (LYMPHOCYTE STIMULATORY FACTOR 1).
 GN IL4.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BEAGLE;
 RX MEDLINE=99105936; PubMed=9887351;
 RA van der Kaaij S.X., Pinelli E., Broeren C.P.M., Schettlers T.P.M.,
 RA Haghighat A., Ruitenberg E.J., Rutten V.P.M.G.;
 RT "Molecular cloning and sequencing of the cDNA for dog interleukin-4.";
 RL Immunogenetics 49:142-143(1999).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN-BEAGLE;
 RA Borouhs K.L., Dreitz M., Sim G.-K.;
 RT "Cloning and expression of canine IL4.";
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE FROM N.A.
 RA Kobayashi H., Yoshida M., Nakagaki K., Katae H., Nogami S.,
 RA Harasawa R., Maeda R., Hayashi Y., Yamamoto H.;
 RT "Nucleotide sequence and deduced amino acid sequence of canine
 interleukin 4 cDNA.";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: IL-4 PARTICIPATES IN AT LEAST SEVERAL B-CELL ACTIVATION
 CC PROCESSES AS WELL AS OF OTHER CELL TYPES. IT IS A COSTIMULATOR OF
 CC DNA-SYNTHESIS. IT INDUCES THE EXPRESSION OF CLASS II MHC MOLECULES
 CC ON RESTING B-CELLS. IT ENHANCES BOTH SECRETION AND CELL SURFACE
 CC EXPRESSION OF IGE AND IGG1. IT ALSO REGULATES THE EXPRESSION OF
 CC THE LOW AFFINITY FC RECEPTOR FOR IGE (CD23) ON BOTH LYMPHOCYTES
 CC AND MONOCYTES (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- SIMILARITY: BELONGS TO THE IL-4 / IL-13 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF054833; AAD11563.1; -
 CC EMBL; AF083270; AAD09005.1; -
 CC EMBL; AF104245; AAC83931.1; -
 CC HSSP; P05112; 1CYL.
 CC InterPro: IPR002354; -
 CC Pfam: PF00727; IL4; 1.
 CC DR PRINTS; PR00431; INTERLEUKIN_4; 1.
 CC DR PROSITE; PS00838; INTERLEUKIN_4_13; 1.
 KW Cytokine; Glycoprotein; B-cell activation; Growth factor; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 132 BY SIMILARITY.
 FT DISULFID 48 84 INTERLEUKIN-4.
 FT DISULFID 70 104 BY SIMILARITY.
 FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 45 45 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 132 AA; 15267 MW; 6F75D8FBB6818945 CRC64;
 Query Match 58.2%; Score 39; DB 1; Length 132;
 Best Local Similarity 58.3%; Pred. No. 8.2;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 EFLERFKSLQK 12
 Db 116 DFLERLAVIMOK 127
 RESULT 6
 ID ATNB_PIG STANDARD; PRT; 303 AA.
 AC P05027;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE 01-OCT-2000 (Rel. 40, Last annotation update)
 DE SODIUM/POTASSIUM-TRANSPORTING ATPASE BETA-1 CHAIN (SODIUM/POTASSIUM-
 DE DEPENDENT ATPASE BETA-1 SUBUNIT).
 GN ATP1B1.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=86220813; PubMed=2423371;
 RA Ovchinnikov Y.A., Modyanov N.N., Broude N.E., Petrukhin K.E.,

RA Grishin A.V., Arzamazova N.M., Aldanova N.A., Monastyrskaya G.S.,
RA Sverdlov E.D.;
RT "Pig Kidney Na⁺,K⁺-ATPase. Primary structure and spatial
RT organization.";
RL FEBS Lett. 201:237-245(1986).
RN [2]
RP SEQUENCE FROM N.A.
RN TISSUE=Kidney;
RA Ovchinnikov Y.A., Broude N.E., Petrukhin K.E., Grishin A.V.,
RA Kiyatkin N.I., Arzamazova N.M., Gevondyan N.M., Chertova E.N.,
RA Melkov A.M., Smirnov Y.V., Malyshev I.V., Monastyrskaya G.S.,
RA Modyanov N.N.;
RT "Nucleotide sequence of cDNA and primary structure of the beta-
RT subunit of Na⁺,K⁺-ATPase from pig kidneys.";
RL Dokl. Biochem. 287:149-152(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=87184715; PubMed=2436627;
RA Broude N.E., Monastyrskaya G.S., Petrukhin K.E., Grishin A.V.,
RA Kiyatkin N.I., Melkov A.M., Smirnov Y.V., Sverdlov V.E.,
RA Malyshev I.V., Modyanov N.N.;
RT "Primary structure of the beta-subunit of Na⁺,K⁺-ATPase from the
RT swine kidney. II. Reverse transcription, cloning of mRNA, complete
RT nucleotide sequence corresponding to the structural region of the
RT gene.";
RL Bioorg. Khim. 13:14-19(1987).
CC -!- FUNCTION: THIS IS THE NON-CATALYTIC COMPONENT OF THE ACTIVE
CC ENZYME, WHICH CATALYZES THE HYDROLYSIS OF ATP COUPLED WITH THE
CC EXCHANGE OF NA AND K IONS ACROSS THE PLASMA MEMBRANE. THE BETA
CC SUBUNIT REGULATES, THROUGH ASSEMBLY OF ALPHA/BETA HETERODIMERS,
CC THE NUMBER OF SODIUM PUMPS TRANSPORTED TO THE PLASMA MEMBRANE.
CC -!- SUBUNIT: COMPOSED OF THREE SUBUNITS: ALPHA (CATALYTIC), BETA
CC AND GAMMA.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE NA⁺/K⁺ AND H⁺ ATPASES BETA CHAIN
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X03937; CAA27575.1; -;
CC EMBL; X04635; CAA28301.1; -;
CC EMBL; M38313; AAA31001.1; -;
CC PIR; A24862; A24862.
CC InterPro: IPR000402; -;
CC Pfam; PF00287; Na_K-ATPase; 1.
CC PROSITE; PS00390; ATPASE_NA_K_BETA_1; 1.
CC PROSITE; PS00391; ATPASE_NA_K_BETA_2; 1.
CC Sodium/potassium transport; Transmembrane; Glycoprotein;
KW Signal-anchor.
FT DOMAIN 1 34 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 35 62 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 63 303 EXTRACELLULAR (POTENTIAL).
FT DISULFID 126 149 BY SIMILARITY.
FT DISULFID 159 175 BY SIMILARITY.
FT DISULFID 213 276 BY SIMILARITY.
FT CARBOHYD 158 178 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CONFLICT 151 16 FI -> LM (IN REF. 3).
FT CONFLICT 151 151 FI -> S (IN REF. 1).
FT CONFLICT 159 159 C -> S (IN REF. 3).
SQ SEQUENCE 303 AA; 35158 MW; 4B9650EDF5942350 CRC64;

Query Match 56.7%; Score 38; DB 1; Length 303;
Best Local Similarity 63.6%; Pred. No. 29;

RA Grishin A.V., Broude N.E., Petrukhin K.E., Grishin A.V.,
RA Kiyatkin N.I., Arzamazova N.M., Gevondyan N.M., Chertova E.N.,
RA Melkov A.M., Smirnov Y.V., Malyshev I.V., Monastyrskaya G.S.,
RA Modyanov N.N.;
RT "Nucleotide sequence of cDNA and primary structure of the beta-
RT subunit of Na⁺,K⁺-ATPase from pig kidneys.";
RL Dokl. Biochem. 287:149-152(1986).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=87184715; PubMed=2436627;
RA Broude N.E., Monastyrskaya G.S., Petrukhin K.E., Grishin A.V.,
RA Kiyatkin N.I., Melkov A.M., Smirnov Y.V., Sverdlov V.E.,
RA Malyshev I.V., Modyanov N.N.;
RT "Primary structure of the beta-subunit of Na⁺,K⁺-ATPase from the
RT swine kidney. II. Reverse transcription, cloning of mRNA, complete
RT nucleotide sequence corresponding to the structural region of the
RT gene.";
RL Bioorg. Khim. 13:14-19(1987).
CC -!- FUNCTION: THIS IS THE NON-CATALYTIC COMPONENT OF THE ACTIVE
CC ENZYME, WHICH CATALYZES THE HYDROLYSIS OF ATP COUPLED WITH THE
CC EXCHANGE OF NA AND K IONS ACROSS THE PLASMA MEMBRANE. THE BETA
CC SUBUNIT REGULATES, THROUGH ASSEMBLY OF ALPHA/BETA HETERODIMERS,
CC THE NUMBER OF SODIUM PUMPS TRANSPORTED TO THE PLASMA MEMBRANE.
CC -!- SUBUNIT: COMPOSED OF THREE SUBUNITS: ALPHA (CATALYTIC), BETA
CC AND GAMMA.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE NA⁺/K⁺ AND H⁺ ATPASES BETA CHAIN
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X03937; CAA27575.1; -;
CC EMBL; X04635; CAA28301.1; -;
CC EMBL; M38313; AAA31001.1; -;
CC PIR; A24862; A24862.
CC InterPro: IPR000402; -;
CC Pfam; PF00287; Na_K-ATPase; 1.
CC PROSITE; PS00390; ATPASE_NA_K_BETA_1; 1.
CC PROSITE; PS00391; ATPASE_NA_K_BETA_2; 1.
CC Sodium/potassium transport; Transmembrane; Glycoprotein;
KW Signal-anchor.
FT DOMAIN 1 34 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 35 62 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 63 303 EXTRACELLULAR (POTENTIAL).
FT DISULFID 126 149 BY SIMILARITY.
FT DISULFID 159 175 BY SIMILARITY.
FT DISULFID 213 276 BY SIMILARITY.
FT CARBOHYD 158 178 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CONFLICT 151 16 FI -> LM (IN REF. 3).
FT CONFLICT 151 151 FI -> S (IN REF. 1).
FT CONFLICT 159 159 C -> S (IN REF. 3).
SQ SEQUENCE 303 AA; 35158 MW; 4B9650EDF5942350 CRC64;

Query Match 56.7%; Score 38; DB 1; Length 303;
Best Local Similarity 63.6%; Pred. No. 29;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 FLERFKSLLOK 12
DB 108 FLEKYKDLAQK 118

RESULT 7
SCPBL_APLSP STANDARD; PRT; 136 AA.
AC P09892;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SMALL CARDIOACTIVE PEPTIDE B PRECURSOR (SCP-B).
OS Aplysia sp. (Sea hare).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Opisthobranchia; Anaspidea;
OC Aplysiidae; Aplysia.
OX NCBI_TaxID=6504;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85216606; PubMed=3858852;
RA Mahon A.C., Lloyd P.E., Weiss K.R., Kupfermann I., Scheller R.H.;
RT "The small cardioactive peptides A and B of Aplysia are derived from
RT a common precursor molecule.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:3925-3929(1985).
CC -!- FUNCTION: THE ACTIVITIES OF SCP-B INCLUDE STIMULATING CONTRACTILE
CC ACTIVITY IN THE GUT, INCREASING THE AMPLITUDE OF THE HEART BEAT,
CC AND ENHANCING THE CONTRACTILE RESPONSE OF THE RADULA CLOSER
CC MUSCLE.
CC -!- MISCELLANEOUS: IN APLYSIA THE HIGHEST CONCENTRATION OF SCP-B IS
CC FOUND IN THE BUCCAL GANGLION.
CC -!- SIMILARITY: AN IDENTICAL REGION TO SCP-A IS FOUND.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M11396; AAA27777.1; -;
KW Neuropeptide; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 136 SMALL CARDIOACTIVE PEPTIDE B.
FT SIMILAR 27 35 TO SCP-A.
SQ SEQUENCE 136 AA; 14773 MW; 8988DB25670CC1EE CRC64;

Query Match 55.2%; Score 37; DB 1; Length 136;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 EFLERFKSLLOK 12
DB 125 EVLKSLKSLLOK 136

RESULT 8
Y875_RICPR STANDARD; PRT; 174 AA.
AC Q92C92;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL PROTEIN RP875.
GN RP875.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria";
RL Nature 396:133-140(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ235273; CAAL5298.1; --
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 7 27 POTENTIAL.
SQ SEQUENCE 174 AA; 20018 MW; D655576742840F3A CRC64;

Query Match 55.2%; Score 37; DB 1; Length 174;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLLOK 12
| :|||:| |
DB 71 ESTERFKLLMK 82

RESULT 9
YP2D.CAEEL STANDARD; PRT; 193 AA.
AC Q09211;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 22.8 KDA PROTEIN AH6.13 IN CHROMOSOME II.
GN AH6.13.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Berks M.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z48009; CAAB8087.1; --
DR WormPep; AH6.13; CE07654.
DR InterPro; IPR000344; --
DR Pfam; PF02117; Str; 2.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 25 45 POTENTIAL.
FT TRANSMEM 107 127 POTENTIAL.
FT TRANSMEM 137 157 POTENTIAL.
SQ SEQUENCE 193 AA; 22779 MW; 09212268B04BAB1 CRC64;

Query Match 55.2%; Score 37; DB 1; Length 193;

Best Local Similarity 58.3%; Pred. No. 26;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EFLERFKSLLOK 12
| :|||:| |
DB 10 EIIERFESIHKQ 21

RESULT 10
Y21C_MYCPN STANDARD; PRT; 357 AA.
ID Y21C_MYCPN
AC P75470; O08089;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN MG218.1 HOMOLOG (F10_ORF357).
GN MPN311 OR MP525.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97103885; PubMed=8948633;
RA Himmelfreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97252497; PubMed=9098066;
RA Krause D.C., Proft T., Hedreyda C.T., Hilbert H., Plagens H.,
RA Herrmann R.;
RT "Transposon mutagenesis reinforces the correlation between Mycoplasma
RT pneumoniae cytoskeletal protein HMW2 and cytoadherence.";
RL J. Bacteriol. 179:2668-2677(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000051; AAB96173.1; --
DR EMBL: U59896; AAB52528.1; --
KW Hypothetical protein.
SQ SEQUENCE 357 AA; 40580 MW; F828443341D3D7F9 CRC64;

Query Match 55.2%; Score 37; DB 1; Length 357;
Best Local Similarity 77.8%; Pred. No. 50;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 ERFKSLLOK 12
| :|||:| |
DB 199 QRFKTLLOK 207

RESULT 11
CP11_MACFA STANDARD; PRT; 512 AA.
ID CP11_MACFA
AC P33616; Q29489;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CYTOCHROME P450 1A1 (EC 1.14.14.1) (CYP1A1).
GN CYP1A1.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92256486; PubMed=1581357;
 RA Komori M., Kikuchi O., Kitada M., Kamatani T.;
 RT "Molecular cloning of monkey P450 1A1 cDNA and expression in yeast.";
 RL Biochem. Biophys. Acta 1131:23-29(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Ohmachi T., Sagami I., Kikuchi H., Fujii H., Suzuki Y., Fujiwara T.,
 RA Watanabe M.;
 RT "Molecular cloning and sequence analysis of cDNA encoding a crab-
 RT eating monkey (Macaca irus) cytochrome P-450";
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
 CC MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
 CC NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
 CC OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
 CC ACIDS, AND XENOBIOTICS.
 CC -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
 CC OXIDIZED FLAVOPROTEIN + H(2)O.
 CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D17575; BAA04500.1; -
 DR PIR; S21761; S21761; -
 DR InterPro; IPR001128; -
 DR InterPro; IPR002401; -
 DR Pfam; PF00067; p450; 1.
 DR PRINTS; PR00385; P450.
 DR PRINTS; PR00463; EP4501.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum.
 FT BINDING 457 457 HEME (BY SIMILARITY).
 FT CONFLICT 277 277 Y -> H (IN REF. 2).
 FT CONFLICT 490 490 M -> V (IN REF. 2).
 SQ SEQUENCE 512 AA; 58155 MW; 239A3237E3870ED7 CRC64;
 Query Match 55.2%; Score 37; DB 1; Length 512;
 Best Local Similarity 63.6%; Pred. No. 73;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 OY 4 ERKSLQKMI 14
 I:|:|:|:|:|
 DB 256 EKFSFMOKMI 266
 RESULT 12
 ID APM2_YEAST STANDARD; PRT; 605 AA.
 AC P38700;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE ADAPTIN MEDIUM CHAIN HOMOLOG APM2.
 GN APM2 OR YHL019C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=S288C / REE526;
 RA MEDLINE=95268148; PubMed=7749194;
 RA Stepp J.D., Pellicena-Palle A., Hamilton S., Kirchhausen T.,
 RA Lemmon S.K.;
 RT "A late Golgi sorting function for Saccharomyces cerevisiae Apm1p,
 RT but not for Apm2p, a second yeast clathrin AP medium chain-related
 RT protein.";
 RL Mol. Biol. Cell 6:41-58(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE=94378003; PubMed=8091229;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 RA Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
 RA Latreille P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
 RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
 RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
 RA Vaudin M.;
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 RT VIII.";
 RL Science 265:2077-2082(1994).
 CC -!- SIMILARITY: BELONGS TO THE ADAPTOR COMPLEXES MEDIUM SUBUNITS
 CC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U09841; AAA83415.1; -
 DR EMBL; U11582; CAB34896.1; -
 DR PIR; S46833; S46833.
 DR SGD; S0001011; APM2.
 DR InterPro; IPR001392; -
 DR Pfam; PF00928; Adap.comp.sub; 1.
 DR PROSITE; PS00990; CLAT_ADAPTOR_M_1; 1.
 DR PROSITE; PS00991; CLAT_ADAPTOR_M_2; 1.
 KW Coated pits.
 SQ SEQUENCE 605 AA; 69990 MW; 7E216B11325EE3C CRC64;
 Query Match 55.2%; Score 37; DB 1; Length 605;
 Best Local Similarity 72.7%; Pred. No. 87;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 2 FLERFKSLQK 12
 I:|:|:|:|:|
 DB 82 FLEQFYHLQK 92
 RESULT 13
 ID EDD_HELPJ STANDARD; PRT; 608 AA.
 AC Q9ZRB3;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PHOSPHOGLUCONATE DEHYDRATASE (EC 4.2.1.12) (6-PHOSPHOGLUCONATE
 DE DEHYDRATASE).
 GN EDD OR JHP1026.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=85963;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tumblino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC
CC -!- CATALYTIC ACTIVITY: 6-PHOSPHO-D-GLUCONATE = 6-PHOSPHO-2-DEHYDRO-
CC 3-DEOXY-D-GLUCONATE + H(2)O.
CC -!- PATHWAY: KEY ENZYME IN THE ENTNER-DOUDOROFF PATHWAY.
CC -!- SIMILARITY: BELONGS TO THE ILVD / EDD FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB001530; AAD06597.1; -.
CC InterPro: IPR000581; -.
CC Pfam: PF00920; ILVD_EDD_1; 1.
CC PROSITE; PS00886; ILVD_EDD_1; 1.
CC PROSITE; PS00887; ILVD_EDD_2; 1.
CC Lyase.
CC SEQUENCE 608 AA; 66603 MW; 978A046F3AE15F98 CRC64;

Query Match 55.2%; Score 37; DB 1; Length 608;
Best Local Similarity 87.5%; Pred. No. 88;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFLERFKS 8
DB 455 EFLERFKN 462

RESULT 14
EDD_HELPY
ID EDD_HELPY STANDARD; PRT; 608 AA.
AC P56111;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PHOSPHOGLUCONATE DEHYDRATASE (EC 4.2.1.12) (6-PHOSPHOGLUCONATE
DE DEHYDRATASE).
GN EDD OR HP1100.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
CC -!- CATALYTIC ACTIVITY: 6-PHOSPHO-D-GLUCONATE = 6-PHOSPHO-2-DEHYDRO-
CC 3-DEOXY-D-GLUCONATE + H(2)O.
CC -!- PATHWAY: KEY ENZYME IN THE ENTNER-DOUDOROFF PATHWAY.

```

```

CC -!- SIMILARITY: BELONGS TO THE ILVD / EDD FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000616; AAD08143.1; -.
CC TIGR; HP1100.
CC InterPro: IPR000581; -.
CC Pfam: PF00920; ILVD_EDD; 1.
CC PROSITE; PS00886; ILVD_EDD_1; 1.
CC PROSITE; PS00887; ILVD_EDD_2; 1.
CC Lyase.
CC SEQUENCE 608 AA; 66655 MW; 47EF7E62E3371F59 CRC64;

Query Match 55.2%; Score 37; DB 1; Length 608;
Best Local Similarity 87.5%; Pred. No. 88;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFLERFKS 8
DB 455 EFLERFKN 462

RESULT 15
YMW6_YEAST
ID YMW6_YEAST STANDARD; PRT; 898 AA.
AC Q04748;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 104.7 KDA PROTEIN IN NCAL-HMS1 INTERGENIC REGION.
GN YMR066W OR YMR916.05.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Pearson D., Bowman S., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z48952; CAA88791.1; -.
CC SGD; S0004670; YMR066W.
CC Hypothetical protein.
CC SEQUENCE 898 AA; 104747 MW; 5707F3ACEEC751F7 CRC64;

Query Match 55.2%; Score 37; DB 1; Length 898;
Best Local Similarity 53.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EFLERFKSLLQKM 13
DB 614 EFKQKVKSLKRM 626

Search completed: May 23, 2001, 11:20:28

```


GenCore version 4.5
Copyright (C) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:19:47 ; Search time 118.04 Seconds
(without alignments)
13.901 Million cell updates/sec

Title: US-09-522-217-2_COPY_135_148

Perfect score: 67

Sequence: 1 EFLERFKSLQKMI 14

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SPTRMBL15.*
- 2: sp-archaea.*
- 3: sp-bacteria.*
- 4: sp-fungi.*
- 5: sp-human.*
- 6: sp-invertebrate.*
- 7: sp-mammal.*
- 8: sp-mhc.*
- 9: sp-organelle.*
- 10: sp-phage.*
- 11: sp-plant.*
- 12: sp-rodent.*
- 13: sp-unclassified.*
- 14: sp-vertebrate.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	62.7	455	5 Q9TZC6	Q9tzc6 caenorhabdi
2	41	61.2	205	2 Q31560	Q31560 bacillus su
3	41	61.2	570	5 Q9XW62	Q9xw62 caenorhabdi
4	41	61.2	2144	5 Q62218	Q62218 caenorhabdi
5	40	59.7	154	10 Q04879	Q04879 arabidopsis
6	40	59.7	305	10 Q9LU64	Q9lu64 arabidopsis
7	40	59.7	506	5 Q18822	Q18822 caenorhabdi
8	39	58.2	329	10 Q43737	Q43737 arabidopsis
9	39	58.2	329	10 Q9LSY7	Q9lsy7 arabidopsis
10	39	58.2	399	5 Q9VNC4	Q9vnc4 drosophila
11	39	58.2	414	4 Q92552	Q92552 homo sapien
12	39	58.2	490	2 Q33965	Q33965 plectonema
13	39	58.2	613	2 Q9PN30	Q9pn30 campylobact
14	39	58.2	1817	5 Q9G206	Q9g206 plasmodium
15	38	56.7	317	14 Q12701	Q12701 virus of se
16	38	56.7	406	2 Q30508	Q30508 pseudomonas
17	38	56.7	505	10 Q91NB2	Q91nb2 arabidopsis
18	38	56.7	620	4 Q9UDW0	Q9udw0 homo sapien
19	38	56.7	640	2 Q9WZB2	Q9wzb2 thermotoga

20	38	56.7	672	14 Q9YMU7	Q9ymu7 lymantria d
21	38	56.7	751	5 P91406	P91406 caenorhabdi
22	38	56.7	979	5 Q9VHM2	Q9vhm2 drosophila
23	38	56.7	1366	10 Q9LMQ6	Q9lmq6 arabidopsis
24	38	56.7	5002	5 Q9VXR3	Q9vxr3 drosophila
25	37	55.2	127	2 Q68732	Q68732 versinia pe
26	37	55.2	338	2 Q92513	Q92513 streptomyce
27	37	55.2	383	2 Q55666	Q55666 synechocyst
28	37	55.2	427	1 Q9UXS3	Q9uxs3 acidianus a
29	37	55.2	444	5 Q9N9J5	Q9n9j5 leishmania
30	37	55.2	464	5 Q18691	Q18691 caenorhabdi
31	37	55.2	481	4 Q9NVX4	Q9nvx4 homo sapien
32	37	55.2	483	2 Q9RPY9	Q9rpy9 listeria mo
33	37	55.2	497	3 Q9P303	Q9p303 ciadosporiu
34	37	55.2	508	2 P74942	P74942 thermus aqu
35	37	55.2	510	5 Q97251	Q97251 plasmodium
36	37	55.2	558	2 Q9WYH2	Q9wyh2 thermotoga
37	37	55.2	590	10 Q82363	Q82363 arabidopsis
38	37	55.2	600	8 Q47881	Q47881 brassica ol
39	37	55.2	619	2 Q9PPC0	Q9ppc0 campylobact
40	37	55.2	830	2 Q9S502	Q9s502 myxococcus
41	37	55.2	854	2 Q9RA63	Q9ra63 thermus aqu
42	37	55.2	858	5 Q9V429	Q9v429 drosophila
43	37	55.2	1288	10 Q9STV0	Q9stv0 arabidopsis
44	37	55.2	1551	5 Q9VMF3	Q9vmf3 drosophila
45	37	55.2	2207	3 Q93845	Q93845 emericella

ALIGNMENTS

RESULT 1

ID	Q9TZC6	PRELIMINARY;	PRT;	455 AA.
AC	Q9TZC6;			
DT	01-MAY-2000	(TREMBLrel. 13, Created)		
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)		
DT	01-OCT-2000	(TREMBLrel. 15, Last annotation update)		
DE	Y59C2A.1	PROTEIN.		
GN	Y59C2A.1.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;			
OC	Rhabditidae; Peloderinae; Caenorhabditis.			
OX	NCBI_TaxID:6239;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BRISTOL N2;			
RX	MEDLINE=94150718; PubMed=7906398;			
RA	Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,			
RA	Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,			
RA	Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,			
RA	Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,			
RA	Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,			
RA	Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,			
RA	Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,			
RA	Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,			
RA	Thierry-Mieg J., Thomas K., Vaubin K., Vaughan K., Waterston R.,			
RT	Watson A., Weinstock L., Wilkison-Sproat J., Wohlman P.;			
RT	"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.			
RT	elegans.";			
RL	Nature 368:32-38(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BRISTOL N2;			
RC	Ozersky P.;			
RT	"The sequence of C. elegans cosmid Y59C2A.";			
RT	Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BRISTOL N2;			
RC	Waterston R.;			
RA	Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF099003; AAC68743.1; -			

DR HSSP: P00730; 5CPA.
 DR INTERPRO: IPR000276; -.
 DR INTERPRO: IPR000834; -.
 DR PFAM: PF00246; Zn_carboxypept; 1.
 DR PRINTS: PR00765; CRBOXYPTASEA.
 DR PROSITE: PS00132; CARBOXYPEPT_ZN_1; UNKNOWN_1.
 DR PROSITE: PS00237; G_PROTEIN_RECEPTOR; UNKNOWN_1.
 SQ SEQUENCE 455 AA; 52072 MW; 51B2EC2EC50C9C4A CRC64;

Query Match 62.7%; Score 42; DB 5; Length 455;
 Best Local Similarity 80.0%; Pred. No. 67;
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EFLERFKSL 10
 |||:|||||
 Db 89 EFLQFKSL 98

RESULT 2

ID O31560 PRELIMINARY; PRT; 205 AA.
 AC O31560;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE YFIR PROTEIN.
 GN YFIR.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;

[1]

SEQUENCE FROM N.A.

STRAIN=168;

MEDLINE=98044033; PubMed=9384377;

Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

Azevedo V., Bertero M.C., Bessieres P., Bolotin A., Borchert S.,

Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,

Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,

Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,

Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,

Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,

Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,

Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,

Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,

Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,

Kobayashi Y., Koetter P., Konigstein G., Krohn S., Kumano M.,

Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,

Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

Parro V., Pohl T.M., Portetelle D., Portwolk S., Prescott A.M.,

Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

Rieger M., Rivolta C., Rochelle E., Roche B., Rose M., Sadale Y.,

Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,

Seikiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,

Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,

Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,

Tosato V., Uchiyama S., Vandenbol M., Vannier P., Vassartotti A.,

Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,

Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,

Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,

"The complete genome sequence of the gram-positive bacterium Bacillus

subtilis";

Nature 390:249-256(1997).

[2]

SEQUENCE FROM N.A.

STRAIN=168;

Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;

Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.

[3]

RP SEQUENCE FROM N.A.
 RX MEDLINE=97101647; PubMed=8946165;
 RY Yamamoto H., Uchiyama S., Sekiguchi J.;
 RT "Cloning and sequencing of a 27.8-kb nucleotide sequence of the 79
 degrees-81 degrees region of the Bacillus subtilis genome containing
 the sspE locus";
 RL DNA Res. 3:257-262(1996).
 DR EMBL; Z99108; CAB12666.1; -.
 DR EMBL; D85082; BAA24458.1; -.
 DR INTERPRO: IPR001647; -.
 DR PIR: P00440; tetr; 1.
 SQ SEQUENCE 205 AA; 23684 MW; 3D075561981F49D0 CRC64;

Query Match 61.2%; Score 41; DB 2; Length 205;
 Best Local Similarity 69.2%; Pred. No. 44;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 2 FLERFKSLQKMI 14
 |::||| |||||
 Db 134 FVERFSRLQKGI 146

RESULT 3

Q9XW62
 ID Q9XW62 PRELIMINARY; PRT; 570 AA.
 AC Q9XW62;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE Y7588A.25 PROTEIN.
 GN Y7588A.25.

OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;

[1]

SEQUENCE FROM N.A.

MEDLINE=94150718; PubMed=7906398;

Barlow K.;

"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

elegans";

Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

[2]

SEQUENCE FROM N.A.

MEDLINE=94150718; PubMed=7906398;

Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,

Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,

Craighton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,

Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,

Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,

Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,

Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,

Smalton N., Smith A., Sonnhammer E., Staden R., Sulston J.,

Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,

Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;

"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

elegans";

Nature 368:32-38(1994).

DR EMBL; AL033514; CAA22109.1; -.

SQ SEQUENCE 570 AA; 64360 MW; 618998C448D3B5F9 CRC64;

Query Match 61.2%; Score 41; DB 5; Length 570;
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 EFLERFKSLQK 12
 |::||| |||||

Db 343 EILEFFKSLQK 354

RESULT 4

```

O62218 ID O62218 PRELIMINARY; PRT; 2144 AA.
AC O62218;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE F33H2.5 PROTEIN.
GN F33H2.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Cottage A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lighning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlmann P.;
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans*;
RL Nature 368:32-38(1994).
DR EMBL; Z81526; CAB04263.1; -.
DR INTERPRO; IPR002064; -.
DR PFAM; PF00136; DNA_pol.B; 3.
DR PRINTS; PRO0106; DNAPOLB.
SQ SEQUENCE 2144 AA; 244706 MW; D73E820585E81B75 CRC64;

Query Match 61.2%; Score 41; DB 5; Length 2144;
Best Local Similarity 61.5%; Pred. No. 4.4e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 FLERFKSLQKMI 14
:|||||:|:|:|
DB 1112 YLRFSGCIQKII 1124

RESULT 5
ID O04879 PRELIMINARY; PRT; 154 AA.
AC O04879;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE SUPEROXIDE DISMUTASE (EC 1.15.1.1) (FRAGMENT).
GN SODB OR SOD.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Van Bruseghem F., Villarroel R., van Montagu M., Inze D.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
DR EMBL; Y12641; CAA73188.1; -.
DR HSP; P09157; IISC.

```

```

DR MENDEL; 16443; Arath; Sodb; 16443.
DR INTERPRO; IPR001189; -.
DR PFAM; PF00081; sodfr; 1.
DR PROSITE; PS00088; SOD_MN; 1.
KW Oxidoreductase.
FT NON_TER 1
SQ SEQUENCE 154 AA; 17802 MW; 862B8D0C5CBFFCAD CRC64;

Query Match 59.7%; Score 40; DB 10; Length 154;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFLERFKS 8
|||||||
DB 10 EFLERFKS 17

RESULT 6
Q9LU64 PRELIMINARY; PRT; 305 AA.
AC Q9LU64;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE SIMILARITY TO IRON-CONTAINING SUPEROXIDE DISMUTASE.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty pl and TAC
RT clones.";
RL DNA Res. 7:31-63(2000).
DR EMBL; AB023044; BAA97372.1; -.
SQ SEQUENCE 305 AA; 34664 MW; 6A68EAF701EA5AC2 CRC64;

Query Match 59.7%; Score 40; DB 10; Length 305;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EFLERFKS 8
|||||||
DB 161 EFLERFKS 168

RESULT 7
ID Q18822 PRELIMINARY; PRT; 506 AA.
AC Q18822;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE SIMILAR TO ALDEHYDE DEHYDROGENASE. NCBI GI: 1166604.
GN C54D1.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,

```

RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P., III of C.
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RL elegans";
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Mixx M.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U46673; AAC48151.1; -;
 DR HSSP; P20000; 1AG8.
 DR INTERPRO; IPR002086; -;
 DR PFAM; PF00171; aldedh; 1.
 DR PROSITE; PS00070; ALDEHYDE-DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE; PS00687; ALDEHYDE-DEHYDR_GLU; UNKNOWN_1.
 SQ SEQUENCE 506 AA; 55388 MW; 266ADF274B8D9829 CRC64;

Query Match 59.7%; Score 40; DB 5; Length 506;
 Best Local Similarity 53.8%; Pred. No. 1.5e+02;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 FLERFKSLQRM 14
 : : : : :
 Db 6 YREEFKNLIQKL 18

RESULT 8
 Q43737 PRELIMINARY; PRT; 329 AA.
 AC Q43737; O96521;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE PEROXIDASE PRECURSOR (EC 1.11.1.7) (MYELOPEROXIDASE) (PEROXIDASE ATP7A
 DE PRECURSOR).
 GN YP99 OR PRXR9
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VARIOUS STRAINS;
 RA Capelli N., Tognolli M., Flach J., Overney S., Penel C., Greppin H.,
 RA Simon P.;
 RL plant Physiol. 112:446-446(1996).
 RN [2]
 RP SEQUENCE OF 4-329 FROM N.A.
 RA Wellinder K.G., Jespersen H.M., Kjaersgaard I.V.H., Justesen A.F.,
 RA Oestergaard L., Abelskov A.K., Jensen R.B., Hansen L.N.,
 RA Rasmussen S.K.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: DONOR + H(2)O(2) = OXIDIZED DONOR + 2 H(2)O.
 CC -1- COFACTOR: HEME.
 DR EMBL; X98321; CAA66965.1; -;
 DR EMBL; X98854; CAA67360.1; -;
 DR HSSP; P00433; 2ATJ.
 DR MENDEL; 6571; Arath.Ypr9;6571.
 DR INTERPRO; IPR000823; -;
 DR INTERPRO; IPR002016; -;
 DR PFAM; PF00141; peroxidase; 1.
 DR PRINTS; PR00458; PEROXIDASE.
 DR PRINTS; PR00461; PEROXIDASE.
 DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
 DR PROSITE; PS00436; PEROXIDASE_2; 1.

KW Signal; Peroxidase; Oxidoreductase.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 329 POTENTIAL.
 SQ SEQUENCE 329 AA; 35815 MW; 693B6941617393A0 CRC64;

Query Match 58.2%; Score 39; DB 10; Length 329;
 Best Local Similarity 58.3%; Pred. No. 1.4e+02;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 FLERFKSLQKM 13
 : : : : :
 Db 104 FVERIKALLEKV 115

RESULT 9
 Q9LSV7 PRELIMINARY; PRT; 329 AA.
 AC Q9LSV7;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE PEROXIDASE.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
 RT clones.";
 RT DNA Res. 7:131-135(2000).
 DR EMBL; AB025634; BAB02839.1; -;
 SQ SEQUENCE 329 AA; 35788 MW; 78CE6941617393A0 CRC64;

Query Match 58.2%; Score 39; DB 10; Length 329;
 Best Local Similarity 58.3%; Pred. No. 1.4e+02;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 FLERFKSLQKM 13
 : : : : :
 Db 104 FVERIKALLEKV 115

RESULT 10
 Q9VNC4 PRELIMINARY; PRT; 399 AA.
 AC Q9VNC4;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE CG12167 PROTEIN.
 GN CG12167.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananadides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon K.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cayley S., Dahke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 DR ENBL; AE003602; AAF52020.1;
 DR FLYBASE; FBgn0037333; CGI2167.
 DR INTERPRO; IPR002965;
 DR PRINTS; PR01217; PRICHEXTENSN.
 SQ SEQUENCE 399 AA; 42870 MW; 39C385009176FA4C CRC64;

Query Match 58.2%; Score 39; DB 5; Length 399;
 Best Local Similarity 63.6%; Pred. No. 1.7e+02;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 FLERFKSLLOK 12
 ||| || :||
 Db 154 FLENFKKILEK 164

RESULT 11
 Q92552 PRELIMINARY; PRT; 414 AA.
 AC Q92552;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE MITOCHONDRIAL 28S RIBOSOMAL PROTEIN S27 (MRP-S27).
 GN KIAA0264.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BONE MARROW;
 RX MEDLINE=97191544; PubMed=9039502;
 RA Nagase T., Seki N., Ishikawa K., Ohira M., Kawarabayashi Y., Ohara O.,
 RA Tanaka A., Kotani H., Miyajima N., Nomura N.;
 RT "Prediction of the coding sequences of unidentified human genes. VI.

RT The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
 RT analysis of cDNA clones from cell line KG-1 and brain.";
 RL DNA RES. 3:321-329(1996).
 RN [2]

RP IDENTIFICATION.
 RA Koc E.C., Burkhardt W., Blackburn K., Moseley A., Koc H.,
 RA Spremulli L.L.;
 RA "Identification of four proteins from the small subunit of the
 RT mammalian mitochondrial ribosome using a proteomics approach.";
 RL Submitted (JUL-2000) to the SWISS-PROT data bank.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
 DE EMBL; D87453; BAA13394.1; ALT INIT.
 KW Ribosomal protein; Mitochondrion.
 SQ SEQUENCE 414 AA; 47669 MW; AC418C70C5E8AA77 CRC64;

Query Match 58.2%; Score 39; DB 4; Length 414;
 Best Local Similarity 53.8%; Pred. No. 1.8e+02;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EFLERFKSLLOKMI 13
 :|||:| |
 Db 329 QYLERFKALHSKL 341

RESULT 12
 O33965 PRELIMINARY; PRT; 490 AA.
 AC O33965;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE KINESIN LIGHT CHAIN.
 GN KLC.
 OS Plectonema boryanum.
 OC Bacteria; Cyanobacteria; Oscillatoriales; Plectonema.
 OX NCBI_TaxID=1184;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UTEX 485;
 RX MEDLINE=97355691; PubMed=9212172;
 RA Celerin M., Gilpin A.A., Dossantos G., Laudenbach D.E., Clarke M.W.,
 RA Beushausen S.;
 RT "Kinesin light chain in a eubacterium.";
 RL DNA Cell Biol. 16:787-795(1997).
 DR EMBL; U78597; AAB87735.1;
 DR INTERPRO; IPR001440;
 DR PFAM; PF00515; TPR; 7;
 SQ SEQUENCE 490 AA; 55967 MW; B82158CFB2BC9814 CRC64;

Query Match 58.2%; Score 39; DB 2; Length 490;
 Best Local Similarity 57.1%; Pred. No. 2.1e+02;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLLOKMI 14
 | :||:| | | |
 Db 451 EASQRFSLLOKAL 464

RESULT 13
 Q9PN30 PRELIMINARY; PRT; 613 AA.
 ID Q9PN30;
 AC Q9PN30;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE HYPOTHETICAL PROTEIN CJ1268C.
 GN CJ1268C.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
 OC Campylobacter.
 OX NCBI_TaxID=197;

```
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=NCTC 11168;
RA MEDLINE=20150912; PubMed=10688204;
RX Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,
RA Jagels K., Karleyshov A., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., Van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
DR EMBL: AL139077; CAB73522.1;
KW Hypothetical protein.
SQ SEQUENCE 613 AA; 71590 MW; 9B84991266818184 CRC64;

Query Match 58.2%; Score 39; DB 2; Length 613;
Best Local Similarity 50.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 EFLERFKSLQKMI 14
   |||||
DB 103 EYEEFKELLEQFL 116

RESULT 14
O96206 PRELIMINARY; PRT; 1817 AA.
AC O96206;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE PREDICTED SECRETED PROTEIN.
GN PFB0565W.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99021743; PubMed=9804551;
RA Gardner M.J., Tettelin H., Carucci D.J., Cummings L.M., Aravind L.,
RA Koonin E.V., Shalom S., Mason T., Yu K., Fujii C., Pederson J.,
RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M.,
RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Science 282:1126-1132(1998).
DR EMBL: AE001403; AAC71902.1;
SQ SEQUENCE 1817 AA; 217996 MW; D1812785960E0BCE CRC64;

Query Match 58.2%; Score 39; DB 5; Length 1817;
Best Local Similarity 57.1%; Pred. No. 7.7e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EFLERFKSLQKMI 14
   |||||
DB 1793 EFLARFKYLDLII 1806

RESULT 15
O12701 PRELIMINARY; PRT; 317 AA.
AC O12701;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-JUL-1997 (TrEMBLrel. 04, Last annotation update)
DE VSH-1 ASSOCIATED PROTEIN 1.
OS virus of Serpulina hyodysenteriae 1.
OC Viruses; dsDNA viruses, no RNA stage; unclassified dsDNA phages.
OX NCBI_TaxID=58620;
```

```
[1]
RN SEQUENCE FROM N.A.
RP Thompson M.G., Stanton T.B.;
RA Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U90539; AAB51754.1;
SQ SEQUENCE 317 AA; 36477 MW; 333C2901C1340E49 CRC64;

Query Match 56.7%; Score 38; DB 14; Length 317;
Best Local Similarity 35.7%; Pred. No. 2e+02;
Matches 5; Conservative 8; Mismatches 1; Indels 0; Gaps 0;

QY 1 EFLERFKSLQKMI 14
   |||||
DB 155 EYMQKLKNVLEKLI 168

Search completed: May 23, 2001, 11:19:48
Job time: 601 sec
```


RESULT	1
B18623	

```

RESULT 2
BI8627
ID BI8627 standard; Protein; 519 AA.
XX
AC BI8627;
XX
XX
DT 22-JAN-2001 (first entry)
XX
DE Amino acid sequence of MBP-human zalphall ligand fusion in pTAP126.
XX
XX zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
KW tumorigenesis; leukaemia; hematopoiesis; B cell tumour.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX WO200053761-A2.
XX
XX 14-SEP-2000.
XX
XX 09-MAR-2000; 2000WO-US06067.
XX
XX 09-MAR-1999; 99US-0264908.
PR 11-MAR-1999; 99US-0265992.
PR 01-JUL-1999; 99US-0142013.
XX
XX (ZYMO ) ZYMOGENETICS INC.
PA
XX Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX
XX WPI; 2000-565600/52.
DR N-PSDB; A75599.
XX
XX New human cytokine, designated zalphall ligand, useful for stimulating
PT the proliferation and/or development of haematopoietic cells in vitro
PT and in vivo, and for treating tumorigenesis.
XX
XX Example 31; Page 233-235; 256pp; English.
XX
XX The present sequence represents a MBP-human zalphall ligand fusion in
CC the plasmid pTAP126. Zalphall ligand is a cytokine. The zalphall ligand
CC is useful for stimulating the proliferation and development of
CC haematopoietic cells in vitro and in vivo. Zalphall ligand
CC polynucleotides can be used as primers or probes for cloning the
CC zalphall gene. The zalphall ligand is useful for treating tumorigenesis.
CC A zalphall ligand-saporin fusion toxin may be used for treating
CC leukaemias and lymphomas. Antagonists against zalphall ligand are
CC useful as research reagents for characterizing ligand-receptor
CC interaction. Antagonists are also useful for inhibiting expansion,
CC proliferation, activation and differentiation of cells involved in
CC regulating hematopoiesis. The zalphall ligand may also be used to
CC stimulate an immune response against B cell tumour, a virus, a parasite
CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
CC agonists and antibodies are also useful for the detection, diagnosis,
CC prevention, and treatment of diseases associated with a zalphall ligand
CC genetic defect.
XX
XX Sequence 519 AA;
SQ

Query Match 80.98; Score 131; DB 21; Length 519;
Best Local Similarity 100.0%; Pred. No. 8.7e-125;
Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 QDRHMIRMQLDIVDQKLVNVDLVPEFLPAPDEVETNCWSAFSCFOKQKLSANTGN 91
DB 389 qdrhmirmrqlidivdqklnyvndlvpeflpapedevencwsafscfqqklsantgn 448.
OY 92 NERTINVSIKKLRKRPSTNAGRQKRLTCPSDSEYKPKPKFLERFKSLQKMIHQH 151
DB 449 nerlinvsikkkrkppstnagrqrkrltcpsdseyyekppkflerfksllqkmiqh 508

```



```

RESULT 4
BI8626
ID BI8626 standard; Peptide: 32 AA.
XX
AC B18626;
XX
DT 22-JAN-2001 (first entry)
XX
DE AntigenInc peptide derived from a human zalphall ligand polypeptide.
XX
KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
KW tumorigenesis; leukaemia; hematopoiesis; B cell tumour.
XX
OS Homo sapiens.
XX
PN WO200053761-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-US06067.
XX
PR 09-MAR-1999; 99US-0264908.
PR 11-MAR-1999; 99US-0265992.
PR 01-JUL-1999; 99US-0142013.
XX
PA (Zymo ) ZYMOGENETICS INC.
XX
PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX
DR WPI; 2000-565600/52.
XX
PT New human cytokine, designated zalphall ligand, useful for stimulating
PT the proliferation and/or development of haematopoietic cells in vitro
PT and in vivo, and for treating tumorigenesis -
XX
PS Example 34; Page 227; 256pp; English.
XX
CC The present sequence was used to raise antibodies, and is derived from
CC a human zalphall ligand polypeptide. Zalphall ligand is a cytokine.
CC The zalphall ligand is useful for stimulating the proliferation and
CC development of haematopoietic cells in vitro and in vivo. Zalphall
CC ligand polynucleotides can be used as primers or probes for cloning the
CC zalphall gene. The zalphall ligand is useful for treating tumorigenesis.
CC A zalphall ligand-saporin fusion toxin may be used for treating
CC leukaemias and lymphomas. Antagonists against zalphall ligand are
CC useful as research reagents for characterizing ligand-receptor
CC interaction. Antagonists are also useful for inhibiting expansion,
CC proliferation, activation and differentiation of cells involved in
CC regulating hematopoiesis. The zalphall ligand may also be used to
CC stimulate an immune response against B cell tumour, a virus, a parasite
CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
CC agonists and antibodies are also useful for the detection, diagnosis,
CC prevention, and treatment of diseases associated with a zalphall ligand
CC genetic defect.
XX
SQ Sequence 32 AA;

```

```

Query Match 19.8%; Score 32; DB 21; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.3e-25;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 122 CPSCDSYEKKPKPELFERKSLLOKMIHOHLS 153
Db 1 cpscdsyekpkpeflierfkllqkmiqhls 32

```

```

Search completed: May 23, 2001, 11:49:59
Job time: 336 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:47:08 ; Search time 11.37 Seconds
(without alignments)
273.717 Million cell updates/sec

Title: US-09-522-217-2
Perfect score: 162
Sequence: 1 MRSSPGNWERIVICLMVIFL.....LLQKMIHQHLSRTHGSEDS 162

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 185757 seqs, 19210857 residues

Word size : 14

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents.AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description

No matches found

Search completed: May 23, 2001, 11:50:14
Job time: 186 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:47:23 ; Search time 14.38 Seconds
(without alignments)
774.208 Million cell updates/sec

Title: US-09-522-217-2
Perfect score: 162
Sequence: 1 MRSSPGNWERIVICLWIFL.....LLQKWIHQHLSRTHGSEDS 162

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 198801 seqs, 68722935 residues

Word size : 14

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR_67:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	%	
No.	Score	Match Length DB ID	Description

No matches found

Search completed: May 23, 2001, 11:50:32
Job time: 189 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:50:18 ; Search time 9.67 Seconds
(without alignments)
573.877 Million cell updates/sec

Title: US-09-522-217-2
Perfect score: 162
Sequence: 1 MRSSPGNWERIVICLWVFL.....LLQKMIHQHLSRTHGSEDS 162

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 14

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
---------------	-------	----------------	--------	----	----	-------------

No matches found

Search completed: May 23, 2001, 11:52:04
Job time: 106 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:50:03 ; Search time 22.46 Seconds
(without alignments)
845.400 Million cell updates/sec

Title: US-09-522-217-2
Perfect score: 162
Sequence: 1 MRSSPGNWERIVICLWIFL.....LLQKMIHQHLSRTHGSEDS 162

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 374700 seqs, 117207915 residues

Word size : 14

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SPTREMBL_15.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_unclassified.*
13: sp_vertebrate.*
14: sp_virus.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
---------------	-------	----------------	--------	----	----	-------------

No matches found

Search completed: May 23, 2001, 11:51:51
Job time: 108 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:23:23 ; Search time 63.83 Seconds
(without alignments)
130.751 Million cell updates/sec

Title: US-09-522-217-56

Perfect score: 769

Sequence: 1 MERTVCLVIFLGTVAHKS.....KEFLERLKLQKIMQHLS 146

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	769	100.0	146	21	B18624
2	651.5	84.7	510	21	B18628
3	486	63.2	162	21	B18623
4	396.5	51.6	519	21	B18627
5	152	19.8	32	21	B18626
6	126	16.4	40	21	B18625
7	80	10.4	1081	20	Z28482
8	80	10.4	1081	20	W81600
9	75	9.8	632	17	R85300
10	75	9.8	1269	17	W03659
11	74	9.6	670	21	B42416

12	73.5	9.6	871	21	B00191	Breast cancer prot
13	72.5	9.4	321	11	B44987	Human secreted pro
14	71.5	9.3	487	16	R71909	Human histamine H1
15	70.5	9.2	330	16	R79967	Fifth transmembran
16	70.5	9.2	487	16	R79965	Human histamine H1
17	70.5	9.2	874	19	W98698	H. pylori GPO 686
18	69.5	9.0	284	20	Y41211	H. influenzae Yig
19	69	9.0	651	21	Y82492	BYDV coat protein
20	69	9.0	724	21	R82493	BYDV coat protein
21	68.5	8.9	114	16	R83309	Simian interleukin
22	68.5	8.9	114	17	W09101	Human mature epith
23	68.5	8.9	114	19	W39188	Human epithelium d
24	68.5	8.9	114	20	Y03759	Human epithelium d
25	68.5	8.9	114	21	Y52311	Mature human epith
26	68.5	8.9	122	17	R90842	Recombinant flag s
27	68.5	8.9	135	21	Y54825	Human Interleukin-
28	68.5	8.9	162	16	R83436	Simian Interleukin
29	68.5	8.9	162	16	R69226	Simian IL-15. Cer
30	68.5	8.9	162	17	W09100	Human epithelium d
31	68.5	8.9	162	17	W07254	Simian epithelium-
32	68.5	8.9	162	17	R98526	Simian interleukin
33	68.5	8.9	162	17	R92798	Mammalian interleu
34	68.5	8.9	162	19	W39187	Human epithelium d
35	68.5	8.9	162	20	Y03758	Human epithelium-d
36	68.5	8.9	162	21	Y78594	Simian interleukin
37	68.5	8.9	162	21	Y52310	Human epithelium-d
38	68.5	8.9	491	14	R36979	Histamine H1 recep
39	68	8.8	961	20	Y34577	Porphyromonas ging
40	68	8.8	973	20	Y34434	Porphyromonas ging
41	68	8.8	2749	12	R13887	Inositol-3-phospha
42	68	8.8	2749	21	Y77847	Mouse IP3 receptor
43	67.5	8.8	449	21	B36885	S. pneumoniae YycG
44	67	8.7	108	21	Y58205	Canine mature inte
45	67	8.7	132	21	Y58203	Canine interleukin

ALIGNMENTS

RESULT	1
B18624	
ID	B18624 standard; Protein; 146 AA.
XX	XX
AC	B18624;
XX	XX
DT	22-JAN-2001 (first entry)
XX	XX
DE	A mouse zalphall ligand polypeptide.
DE	zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
KW	tumorigenesis; leukaemia; hematopoiesis; B cell tumour.
KW	
XX	XX
OS	Mus musculus.
XX	XX
PN	WO200053761-A2.
XX	XX
PD	14-SEP-2000.
XX	XX
PF	09-MAR-2000; 2000WO-US06067.
XX	XX
PR	09-MAR-1999; 99US-0264908.
PR	11-MAR-1999; 99US-0265992.
PR	01-JUL-1999; 99US-0142013.
XX	XX
XX	(ZYMO) ZYMOGENETICS INC.
XX	XX
PI	Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
PI	Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX	XX
DR	WPI; 2000-565600/52.
DR	N-PSDB; A75580.
XX	XX
PT	New human cytokine, designated zalphall ligand, useful for stimulating


```

XX PF 09-MAR-2000; 2000WO-US06067.
XX PF 09-MAR-1999; 99US-0264908.
XX PR 11-MAR-1999; 99US-0265992.
XX PR 01-JUL-1999; 99US-0142013.
XX PA (ZYMO ) ZYMOGENETICS INC.
XX PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
XX PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX DR WPI; 2000-565600/52.
XX PS
XX PT New human cytokine, designated zalphall ligand, useful for stimulating
XX PT the proliferation and/or development of haematopoietic cells in vitro
XX PT and in vivo, and for treating tumourigenesis -
XX PS Example 34; Page 227; 256pp; English.
XX CC The present sequence was used to raise antibodies, and is derived from
XX CC a human zalphall ligand polypeptide. Zalphall ligand is a cytokine.
XX CC The zalphall ligand is useful for stimulating the proliferation and
XX CC development of haematopoietic cells in vitro and in vivo. Zalphall
XX CC ligand polynucleotides can be used as primers or probes for cloning the
XX CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
XX CC A zalphall ligand-saporin fusion toxin may be used for treating
XX CC leukaemias and lymphomas. Antagonists against zalphall ligand are
XX CC useful as research reagents for characterizing ligand-receptor
XX CC interaction. Antagonists are also useful for inhibiting expansion,
XX CC proliferation, activation and differentiation of cells involved in
XX CC regulating hematopoiesis. The zalphall ligand may also be used to
XX CC stimulate an immune response against B cell tumour, a virus, a parasite
XX CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
XX CC agonists and antibodies are also useful for the detection, diagnosis,
XX CC prevention, and treatment of diseases associated with a zalphall ligand
XX CC genetic defect.
XX SQ Sequence 32 AA;
XX Query Match 19.8%; Score 152; DB 21; Length 32;
XX Best Local Similarity 87.5%; Pred. No. 7.2e-11;
XX Matches 28; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 115 CPCSDSYEKTPKEFLERLKLWLLQKMIHQHLS 146
DB 1 cpscdsyekppkeflerfksllqkmihqhs 32
RESULT 6
B18625
ID B18625 standard; Peptide; 40 AA.
XX AC B18625;
XX AC
XX DT 22-JAN-2001 (first entry)
XX DE Antigeninc peptide derived from a human zalphall ligand polypeptide.
XX KW zalphall ligand; cytokine; haematopoietic cell proliferation; lymphoma;
XX KW tumourigenesis; leukaemia; hematopoesis; B cell tumour.
XX OS Homo sapiens.
XX PN W0200053761-A2.
XX PD 14-SEP-2000.
XX PF 09-MAR-2000; 2000WO-US06067.
XX PR 09-MAR-1999; 99US-0264908.
XX PR 11-MAR-1999; 99US-0265992.

```

```

PR 01-JUL-1999; 99US-0142013.
XX PA (ZYMO ) ZYMOGENETICS INC.
XX PI Novak JE, Presnell SR, Sprecher CA, Foster DC, Holly RD, Gross JA;
XX PI Johnston JV, Nelson AJ, Dillon SR, Hammond AK;
XX DR WPI; 2000-565600/52.
XX PS
XX PT New human cytokine, designated zalphall ligand, useful for stimulating
XX PT the proliferation and/or development of haematopoietic cells in vitro
XX PT and in vivo, and for treating tumourigenesis -
XX PS Example 34; Page 227; 256pp; English.
XX CC The present sequence was used to raise antibodies, and is derived from
XX CC a human zalphall ligand polypeptide. Zalphall ligand is a cytokine.
XX CC The zalphall ligand is useful for stimulating the proliferation and
XX CC development of haematopoietic cells in vitro and in vivo. Zalphall
XX CC ligand polynucleotides can be used as primers or probes for cloning the
XX CC zalphall gene. The zalphall ligand is useful for treating tumourigenesis.
XX CC A zalphall ligand-saporin fusion toxin may be used for treating
XX CC leukaemias and lymphomas. Antagonists against zalphall ligand are
XX CC useful as research reagents for characterizing ligand-receptor
XX CC interaction. Antagonists are also useful for inhibiting expansion,
XX CC proliferation, activation and differentiation of cells involved in
XX CC regulating hematopoiesis. The zalphall ligand may also be used to
XX CC stimulate an immune response against B cell tumour, a virus, a parasite
XX CC or a bacterium. The zalphall polypeptides, polynucleotides, antagonists,
XX CC agonists and antibodies are also useful for the detection, diagnosis,
XX CC prevention, and treatment of diseases associated with a zalphall ligand
XX CC genetic defect.
XX SQ Sequence 40 AA;
XX Query Match 16.4%; Score 126; DB 21; Length 40;
XX Best Local Similarity 64.1%; Pred. No. 1.1e-07;
XX Matches 25; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
QY 26 DRLRLRLHLDIVQLKIYENDLPDPELLSAPQDVKGHC 64
DB 2 drhmirmqlidvqdknyvndlypeflpapedvetnc 40
RESULT 7
Y28482
ID Y28482 standard; Protein; 1081 AA.
XX AC Y28482;
XX AC
XX DT 12-OCT-1999 (first entry)
XX DE cos-1 histidine kinase.
XX DE cos-1; osmosensing histidine kinase; fungal cell wall;
XX KW antifungal compound.
XX OS Candida albicans.
XX PN US5939306-A.
XX PD 17-AUG-1999.
XX PF 16-APR-1997; 97US-0843530.
XX PR 16-APR-1997; 97US-0843530.
XX PA (CALY ) CALIFORNIA INST OF TECHNOLOGY.
XX PA (UYTE-) UNIV TECHNOLOGY CORP.
XX PI Aghan J, Alex LA, Selitrennikoff C, Simon MI;
XX PI

```

DR WPI: 1999-468407/39.
DR N-PSDB; Z00021.
XX
PT Osmosensing histidine kinases, useful for screening anti-fungal
PT compounds for their activity against fungal organisms and for the
PT development of antifungal compounds
XX
PS Disclosure: Column 105-117; 110pp; English.
XX
CC The sequence is the osmosensing histidine kinase cos-1. The sequence is
CC encoded in the DNA sequence Z00021. The Cos-1 protein shows 60% identity
CC and 70% similarity to the N. crassa osmosensing histidine kinase osl.
CC Fungal histidine kinases are involved in the critical adjustment of the
CC fungal cell wall in response to changes in extracellular osmolarity. The
CC histidine kinases are useful for screening anti-fungal compounds for
CC their activity against fungal organisms, as well as for the development
CC of antifungal compounds.
XX
SQ Sequence 1081 AA;

Query Match 10.4%; Score 80; DB 20; Length 1081;
Best Local Similarity 28.5%; Pred. No. 1.9;
Matches 35; Conservative 19; Mismatches 49; Indels 20; Gaps 7;

QY 27 RLRLRLRLHLDIVEQ-LKIYE-----NDLDELISAPQDVKGHCHEAAFCQAKL--KP 79
Db 891 rillekgghlvevengleayeaikrkydvvlmdvqmpvmg-----gfeatekirqwekk 945

QY 80 SNPGNNKTFIDLVLAQLRRRLPARRGGKKQKHIAKPCSDSYEKRTPKFLERLKWLLOK 139
Db 946 snpidsltfrtpilaltahaml----gdrekslak--gmddyvskplkpk1--lmqtink 997

QY 140 MH 142
Db 998 cih 1000

RESULT 8
W81600
ID W81600 standard; Protein; 1081 AA.
XX
AC W81600;
XX
DT 15-MAR-1999 (first entry)
XX
DE Candida albicans CanIK1 protein involved in phenotypic switching.
XX
KW CanIK1; phenotypic switching; two-component hybrid kinase;
KW phenotypic switching; antifungal; antimycotic; fungicide; assay.
XX
OS Candida albicans.
XX
FH Key Location/Qualifiers
FT Domain 509..520
FT /label= H1
FT Domain 623..628
FT /label= N
FT Domain 661..663
FT /label= G1
FT Domain 690..696
FT /label= G2
FT Domain 923..928
FT /label= D
FT Domain 987..1013
FT /label= H2
XX
PN W09855654-A2.
XX
XX 10-DEC-1998.
PD
XX
PF 05-JUN-1998; 98WO-US11658.
XX

PR 06-JUN-1997; 97US-0048914.
XX
PA (IOWA) UNIV IOWA RES FOUND.
XX
PI Soll DR, Srikantha T;
XX
DR WPI: 1999-045801/04.
DR N-PSDB; V69792.
XX
PT Newly isolated Candida albicans two-component hybrid kinase gene,
PT Canik1, encoding a protein associated with phenotypic switching -
PT useful in the identification of pharmaceutical compounds that
PT inhibit pathogenicity in C. albicans
XX
PS Disclosure: Fig 2a-c; 32pp; English.
XX
CC This is the amino acid for Candida albicans CanIK1 protein, a two
CC component hybrid kinase involved in phenotypic switching. It is
CC encoded by a claimed polynucleotide (see V69791). A claimed method
CC of screening potential pharmaceutical compounds comprises adding
CC various yeast cells exhibiting phenotypic switching (at least some
CC of which contain the Canik1 gene, operably linked to a promoter)
CC to a test substance, and evaluating the effect on gene expression.
CC The method substance upon polynucleotide expression. The method is
CC used to identify pharmaceutical compounds that effectively inhibit
CC the pathogenicity of C. albicans.
XX
SQ Sequence 1081 AA;

Query Match 10.4%; Score 80; DB 20; Length 1081;
Best Local Similarity 28.5%; Pred. No. 1.9;
Matches 35; Conservative 19; Mismatches 49; Indels 20; Gaps 7;

QY 27 RLRLRLRLHLDIVEQ-LKIYE-----NDLDELISAPQDVKGHCHEAAFCQAKL--KP 79
Db 891 rillekgghlvevengleayeaikrkydvvlmdvqmpvmg-----gfeatekirqwekk 945

QY 80 SNPGNNKTFIDLVLAQLRRRLPARRGGKKQKHIAKPCSDSYEKRTPKFLERLKWLLOK 139
Db 946 snpidsltfrtpilaltahaml----gdrekslak--gmddyvskplkpk1--lmqtink 997

QY 140 MH 142
Db 998 cih 1000

RESULT 9
R85300
ID R85300 standard; Protein; 632 AA.
XX
AC R85300;
XX
DT 14-APR-1996 (first entry)
XX
DE Arabidopsis pathogen resistance gene RPP5 protein.
XX
KW Pathogen resistant; RPP5; tomato; C. fulvum; Avr 4; Avr 9; fungal;
KW leaf mould; variegation.
XX
OS Arabidopsis sp.
XX
PN W09531564-A2.
XX
PD 23-NOV-1995.
XX
PF 11-MAY-1995; 95WO-GB01075.
XX
PR 07-APR-1995; 95GB-0007232.
PR 11-MAY-1994; 94GB-0009394.
PR 23-DEC-1994; 94WO-GB02812.
PR 31-MAR-1995; 95GB-0006658.
XX

OS Homo sapiens.
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 XX 31-MAR-2000; 2000WO-US08621.
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 XX Shimkets RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR N-PSDB; C76625.
 XX
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 3556-3558; 5507pp; English.
 XX
 CC C74446 to C77606 encode the proteins given in B40237 to B43397, which
 CC represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulneryary;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 670 AA;
 Query Match 9.6%; Score 74; DB 21; Length 670;
 Best Local Similarity 27.3%; Pred. No. 5.2; Mismatches 45; Indels 34; Gaps 6;
 Matches 35; Conservative 14;
 QY 28 LLRLRLHLDIVQLKIYENDLPE-----LLSAPQDVKGHCHEAFAFCQKAKLPSN 81
 Db 184 lfihcrgcrvagsgkpfifevdprgpcftiklsarkdhegsdcyprc-----pnn 236
 QY 82 PGNNKFIIDLVAQLRRRLPARGGKKQKHIAKCP-----SC-----DSYEKRPKPFLE 131
 Db 237 pscppllrmnalehl-----kecehi-kcphskygctgfnqdttyethletcrfe 285
 QY 132 RLAKWLQK 139
 Db 286 glkefiqq 293
 RESULT 12
 ID B00191
 XX B00191 standard; Protein; 871 AA.
 AC B00191;
 XX

DT 08-FEB-2001 (first entry)
 XX
 DE Breast cancer protein BCN5.
 XX
 KW Breast cancer; diagnosis; prognosis; detection; screening;
 KW antibody; oestrogen receptor; anti-oestrogen; immune response;
 KW lymph node; metastases; tumour; BCR3; BCQ8; BCQ5; BCH1; BCN1; BCN2;
 KW BCN5; BCQ2; BCX2; BCX3; BCA2; BCR2; BCJ7; BCY3; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200055629-A2.
 XX
 PD 21-SEP-2000.
 XX
 XX 15-MAR-2000; 2000WO-US06952.
 XX
 PR 15-MAR-1999; 99US-0268865.
 PR 12-NOV-1999; 99US-0439878.
 PR 15-NOV-1999; 99US-0440370.
 PR 16-NOV-1999; 99US-0440493.
 PR 16-NOV-1999; 99US-0440676.
 PR 16-NOV-1999; 99US-0440677.
 PR 29-NOV-1999; 99US-0450810.
 PR 02-DEC-1999; 99US-0453137.
 PR 08-MAR-2000; 2000US-0453137.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Mack D, Gish KC;
 XX
 DR WPI; 2000-638216/61.
 DR N-PSDB; A54124.
 XX
 XX Screening drug candidates for their ability to modulate breast cancer
 PT by contacting the drug to a cell expressing an expression profile gene
 PT and determining modulation of expression of the gene
 XX
 PS Disclosure; Fig 50; 258pp; English.
 XX
 CC New methods for screening drug candidates are described which
 CC comprise adding a drug candidate to a cell that expresses a protein
 CC selected from BCH1, BCA2, BCJ7, BCN1, BCN5, BCQ2, BCQ5, BCR2, BCX2
 CC and BCY3 or their fragments and determining the effect of the drug
 CC on the expression of those proteins. Antibodies to breast cancer
 CC genes (specifically BCH1 or its fragment (BCH1p1 or BCH1p2)) are
 CC useful for inhibiting and treating breast cancer in individuals who
 CC are non-responsive to anti-oestrogen and positive for oestrogen
 CC receptor. Compositions comprising BCH1 or a nucleic acid encoding
 CC BCH1 are useful for eliciting an immune response in an individual.
 CC The antibodies are also useful for the diagnosis and prognosis of
 CC breast cancer and for screening compositions which modulate the
 CC breast cancer phenotype. The method allows rapid and simple
 CC detection of lymph node metastases.
 XX
 SQ Sequence 871 AA;
 Query Match 9.6%; Score 73.5; DB 21; Length 871;
 Best Local Similarity 26.1%; Pred. No. 8.4;
 Matches 41; Conservative 18; Mismatches 41; Indels 57; Gaps 9;
 QY 14 GTVAHKSSPQGPDRLLRLRLHLLI-----DIVQLKIYENDLDEL-----LSAPQDV 60
 Db 325 glvkgsmpydfd---irvpffirgspvsgsvpqf-vlnidlapliadiagldtppdv 380
 QY 61 KGHCHEAFAFCQKAKLK---PSNPGN-----TTFIDLVAQLRRRLPARGG 106
 Db 381 dg-----ksvklldpekpgnrrftnkkakwrdtflvgrgkflrkeeskni 429
 QY 107 KKQKHIAKPCSDSYEKRPKEFLERLKLWLQKMIHQ 143
 Db 430 qgsnhlpk-----yervkelcqaryq 451

```
RESULT 13
B44987
ID B44987 standard; Protein; 321 AA.
XX
AC B44987;
XX
DT 12-FEB-2001 (first entry)
XX
DE Human secreted protein encoded by gene 19 homologue.
XX
KW Secreted protein; human; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective;
KW neotropic; neuroprotective; antibacterial; virucide; fungicide;
KW ophthalmological; gene therapy; treatment; autoimmune disease; infection;
KW cardiovascular disorder; nervous system disorder; ocular disorder;
KW wound healing; epithelial cell proliferation; skin aging; mental state;
KW transplantation; metabolism modulation.
XX
OS Homo sapiens.
XX
PN WO200055200-A1.
XX
PD 21-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-US06042.
XX
PR 12-MAR-1999; 99US-0124143.
PR 03-DEC-1999; 99US-0168663.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
XX
WI; 2000-656008/63.
XX
PT Isolated human secretory proteins, nucleic acids encoding them and
PT antibodies directed against them, useful for diagnosing and treating
PT disorders related to the proteins such as cancer, Alzheimer's disease
PT and Parkinsons -
XX
PS Disclosure; Page 422-423; 453pp; English.
XX
CC This invention describes a novel isolated polypeptide (I) and its
CC encoding nucleic acid molecule (II) which have immunosuppressive,
CC antiarthritic, antirheumatic, antiproliferative, cytostatic, cardiac,
CC vasotropic, cerebroprotective, neotropic, neuroprotective, antibacterial,
CC virucide, fungicide and ophthalmological activity and which can be used
CC for gene therapy. (I) and (II) are used to prevent, treat or ameliorate
CC a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,
CC dogs, chickens or sheep. (I) and (II) are also used in diagnosing a
CC pathological condition or susceptibility to a pathological condition. The
CC antibodies to (II) can also be used in alleviating symptoms associated
CC with the disorders and in diagnostic immunoassays e.g. radioimmunoassays
CC or enzyme linked immunosorbent assays (ELISA). Disorders which are
CC diagnosed or treated include autoimmune diseases e.g. rheumatoid
CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
CC disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders
CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and
CC fungi and ocular disorders e.g. corneal infection. The polypeptides can
CC also be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides are used to
CC modulate mammalian metabolism, to change mammal's mental state or
CC physical state by influencing biorhythms circadian rhythms, depression
CC tendency for violence tolerance for pain, reproductive capabilities,
CC hormonal or endocrine levels, appetite, libido, memory, stress or other
CC cognitive qualities, as a food additive or preservative, such as to
CC increase or decrease storage capabilities, fat content, lipid, protein,
CC carbohydrates, vitamins, minerals, cofactors or other nutritional
```

```
CC components.
XX
SQ Sequence 321 AA;
XX
Query Match 9.4%; Score 72.5; DB 21; Length 321;
Best Local Similarity 26.8%; Pred. No. 3;
Matches 26; Conservative 12; Mismatches 32; Indels 27; Gaps 5;
QY 57 PODVKGH-----CEHAFAFCQAKLK---PSNPGNNKTFIIDLVAO----- 95
DB 211 pselkkmthtgekyqcyfcrcadgsnlkthkshgnnlpyacehpcpafgdere 270
QY 96 LRRRLPARRGGKQKHIAKPCSDSYEKRTPKFLEFLER 132
DB 271 lqrhldlfgghkth---qgphcd---hkstnsddlkr 301
RESULT 14
R71909
ID R71909 standard; Protein; 487 AA.
XX
AC R71909;
XX
DT 29-NOV-1995 (first entry)
XX
DE Human histamine H1 receptor.
XX
KW Human histamine H1 receptor; polymorphism detection; treatments;
KW drug design.
XX
OS Homo sapiens.
XX
PN JP07067654-A.
XX
PD 14-MAR-1995.
XX
PF 03-SEP-1993; 93JP-0219544.
XX
PR 03-SEP-1993; 93JP-0219544.
XX
PA (MITP ) MITSUBISHI YUKA BCL KK.
XX
WI; 1995-143848/19.
XX
DR N-PSDB; Q87915.
XX
PT Human histamine H1 receptor gene and protein - used in a method
PT for the detection of histamine polymorphism
XX
PS Claim 1; Pages 9-12; 13pp; Japanese.
XX
CC Q87915 encodes R71909 the human histamine H1 receptor. The DNA
CC fragment can be used in a new method for the detection of
CC histamine polymorphisms, it may also be used in the development
CC of drugs for the treatment of histamine associated conditions.
XX
SQ Sequence 487 AA;
XX
Query Match 9.3%; Score 71.5; DB 16; Length 487;
Best Local Similarity 28.3%; Pred. No. 6.7;
Matches 26; Conservative 13; Mismatches 26; Indels 27; Gaps 5;
QY 43 KIYENDLPELLSAPODVKGHCEH-----AAFACFQAKLKPSNP-GNNKTFIIDLVAQL 96
DB 212 klykg-----vrhcqhrelnrsfseiklrpenkgdakppgkspew 259
QY 97 RRLPARRGGKQKHIAKPCSDSYEKRTPKPE 128
DB 260 lkrkpkdagggs---vlsps-----qtpke 282
RESULT 15
```

R79967
ID R79967 standard; Protein; 330 AA.
XX
AC R79967;
XX
DT 11-JAN-1996 (first entry)
XX
DE Fifth transmembrane region of histamine H1 receptor.
XX
KW Histamine H1 receptor; human; bovine; transformation; drug screening;
KW pNIV3604B; calcium ion mobilisation; inositol 1,4,5-triphosphate;
KW CHO cell; guanine nucleotide regulatory protein; G protein;
KW hybridisation.
XX
OS Homo sapiens.
XX
PN GB2283239-A.
XX
PD 03-MAY-1995.
XX
PF 29-OCT-1993; 93GB-0022353.
XX
PR 29-OCT-1993; 93GB-0022353.
XX
PA (UNIO) UCB SA.
XX
PI Bollen A, Gillard M, Moguilevsky N, Noyer M, Varsalona F;
XX WPI; 1995-157301/21.
DR N-PSDB; T04235.
XX
PT New nucleic acid encoding human histamine H1 receptor - useful
PT diagnostically and for screening receptor binding drugs
XX
PS Example 1; Page 26-29; 49pp; English.
XX
CC This sequence represents the fifth transmembrane region of histamine H1
CC receptor. This sequence was isolated after screening a lambda gt11 human
CC lung cDNA library with the sequence shown in T04234. The human histamine
CC H1 receptor cDNA sequence was constructed using the sequence
CC encoding this protein, and those shown in T04234 and T04236. Histamine
CC H1 receptor is part of a family of histamine receptor proteins. Of these
CC sequences the greatest homology was found with bovine histamine H1
CC receptor (overall identity of 82%, with the transmembrane regions alone
CC having an identity of 96%). These receptors are coupled to different
CC second messenger pathways via guanine nucleotide regulatory proteins
CC (G proteins). This receptor (human histamine H1) transduces the signal
CC through calcium ion mobilisation via an increase in intracellular
CC inositol 1,4,5-triphosphate levels. Plasmids containing human histamine
CC H1 receptor cDNA (pref. pNIV3604B) are used to transform mammalian cells
CC (pref. CHO cells). These transformed cells express the receptor on the
CC cell surface. The transformed cells can then be used to identify ligands
CC that bind to the encoded protein, esp. for drug screening.
XX
SQ Sequence 330 AA;

Query Match 9.2%; Score 70.5; DB 16; Length 330;
Best Local Similarity 30.7%; Pred. No. 5.3;
Matches 23; Conservative 12; Mismatches 25; Indels 15; Gaps 4;

QY 60 VKGHCEH-----AFAFCKFKALKPSNP-GNNKFTFIDLVAQLRRRLPARRGGKKQHIA 113
|: ||:| :|:| | | | | | : : : | | | |
Db 67 vrqhqdnrlrlnrslpsfseilrpenpkgdakpkdespwelkrkpkdaggs---v1 123

QY 114 KCPSCDSYEKRTPK 128
| | | |
Db 124 ksp-----qtpke 132

Search completed: May 23, 2001, 11:42:44
Job time: 1161 sec

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	80	10.4	1081	2	US-08-843-530B-18	Sequence 18, Appl
2	80	10.4	1117	2	US-08-843-530B-33	Sequence 33, Appl
3	69.5	9.0	284	3	US-09-051-197A-2	Sequence 2, Appl
4	68.5	8.9	114	1	US-08-031-399-3	Sequence 3, Appl
5	68.5	8.9	114	1	US-08-393-305-6	Sequence 6, Appl
6	68.5	8.9	114	1	US-08-726-817-6	Sequence 6, Appl
7	68.5	8.9	114	1	US-08-504-042-3	Sequence 3, Appl
8	68.5	8.9	114	2	US-08-725-969-6	Sequence 6, Appl
9	68.5	8.9	114	2	US-08-794-524-6	Sequence 6, Appl
10	68.5	8.9	114	4	US-09-189-193-6	Sequence 6, Appl
11	68.5	8.9	114	5	PCT-US94-03793-3	Sequence 3, Appl
12	68.5	8.9	122	1	US-08-300-903A-3	Sequence 3, Appl
13	68.5	8.9	162	1	US-08-031-399-2	Sequence 2, Appl
14	68.5	8.9	162	1	US-08-393-305-5	Sequence 5, Appl
15	68.5	8.9	162	1	US-08-284-393B-9	Sequence 9, Appl
16	68.5	8.9	162	1	US-08-726-817-5	Sequence 5, Appl
17	68.5	8.9	162	1	US-08-504-042-2	Sequence 2, Appl
18	68.5	8.9	162	2	US-08-725-969-5	Sequence 5, Appl
19	68.5	8.9	162	2	US-08-794-524-5	Sequence 5, Appl
20	68.5	8.9	162	4	US-09-189-193-5	Sequence 5, Appl
21	68.5	8.9	162	5	PCT-US94-03793-2	Sequence 2, Appl
22	68.5	8.9	162	5	PCT-US95-08950-9	Sequence 9, Appl
23	67	8.7	406	1	US-08-650-275-2	Sequence 2, Appl
24	67	8.7	406	3	US-09-181-318-2	Sequence 2, Appl
25	65.5	8.5	872	1	US-08-451-715A-8	Sequence 8, Appl
26	65	8.5	298	1	US-08-348-792-4	Sequence 4, Appl
27	65	8.5	298	2	US-08-462-738-4	Sequence 4, Appl

QY 27 RLLRLRLDIVEQ-LKIYE-NDLDPELLSAPQDVKGHCHEAAAFACQKAKL--KP 79


```
Db 891 RILEKOGHLEVVEVNGLEAYEATKRKNYDVVLMVQMPVNG-----GFEATEKIQWEKK 945
QY 80 SNPGNNKTFIIDLVAQLRRRLPARRGGKKOKHIAKPCSDSYEKRTPEFLERLKWLLQK 139
Db 946 SNPIDSITFTPIIALTAHAML-----GDRKSLAK--GMDDYVSKPLKPKL--LMQTINK 997
QY 140 MIH 142
Db 998 CIH 1000

RESULT 2
US-08-843-530B-33
; Sequence 33, Application US/08843530B
; Patent No. 5939306
; GENERAL INFORMATION:
; APPLICANT: Selitrennikoff, Claude
; APPLICANT: Agnan, Jacqueline
; APPLICANT: Alex, Lisa A.
; APPLICANT: Simon, Melvin I.
; TITLE OF INVENTION: Osmosensing Histidine Kinases
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/843,530B
; FILING DATE: 16-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MacKnight, Kamrin T.
; REGISTRATION NUMBER: 38,230
; REFERENCE/DOCKET NUMBER: UTC-02717
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1117 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
US-08-843-530B-33

Query Match 10.4%; Score 80; DB 2; Length 1117;
Best Local Similarity 28.5%; Pred. No. 0.52;
Matches 35; Conservative 19; Mismatches 49; Indels 20; Gaps 7;

QY 27 RLIRLRHLIDIVEQ-LKIVE-----NDLDPILLSAPQDVKGHCHEAAAFACFOAKL--KP 79
Db 927 RILEKOGHLEVVEVNGLEAYEATKRKNYDVVLMVQMPVNG-----GFEATEKIQWEKK 981
QY 80 SNPGNNKTFIIDLVAQLRRRLPARRGGKKOKHIAKPCSDSYEKRTPEFLERLKWLLQK 139
Db 982 SNPIDSITFTPIIALTAHAML-----GDRKSLAK--GMDDYVSKPLKPKL--LMQTINK 1033
QY 140 MIH 142
Db 1034 CIH 1036

RESULT 3
```

```
US-09-053-197A-2
; Sequence 2, Application US/09053197A
; Patent No. 6022952
; GENERAL INFORMATION:
; APPLICANT: Weiner, Joel H.
; APPLICANT: Turner, Raymond J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PROTEIN
; SECRETION
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,197A
; FILING DATE: 01-APR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MacKnight, Kamrin T.
; REGISTRATION NUMBER: 38,230
; REFERENCE/DOCKET NUMBER: UALB-03293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 284 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-053-197A-2

Query Match 9.0%; Score 69.5; DB 3; Length 284;
Best Local Similarity 23.3%; Pred. No. 1.6;
Matches 42; Conservative 29; Mismatches 52; Indels 57; Gaps 10;

QY 7 CLIVIF-----LGTVAHKSSPOGPDRLRLRLRLIDIV-----EQLK 43
Db 99 CLILVFQNLFYXWVLGVV-----LGPRLPLTAIRTVMDVWKTIRGLAANVQNELKQELK 153
QY 44 IVE-----NDLDPILLSAPQDVKGHCHEAAAFACFOAKL--KPSNPGNNKTF 88
Db 154 LQELQDSIKKAESLNQALSPELSKTVEELKAQADK-----MKAELDKAAQAG---TT 204
QY 89 IIDLVAQLRRRLPARRGGKKOKHIAKPCSD--SYEKRTPEFLERLKWLLQKMIHQLS 146
Db 205 VEDQIKEIKS--AAENAEKSNQNAISVEEAETLSEAERTPTD-LTALETHEKVELNTHLS 261

RESULT 4
US-08-031-399-3
; Sequence 3, Application US/08031399
; Patent No. 5552303
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eismann, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Epithelium-derived T-cell Factor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
```

STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/031,399
FILING DATE: 19930308
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Launer, Charlene
REGISTRATION NUMBER: 33,035
REFERENCE/DOCKET NUMBER: 2811
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-587-0430
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 114 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO

US-08-031-399-3

```

Query Match      8.9%; Score 68.5; DB 1; Length 114;
Best Local Similarity 22.8%; Pred. No. 0.63;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps 4;

QY 32 LRHLIDIVEQLKIYENDLPPELLSAPOOVKGHCHEAFACQAKLKPSNGNNKFTIID 91
   | : | : : | | | | | | | | | | | : | : | |
Db 9 LKKIEDLIQSMHI-----DALTYESDVHPSCVKMTAMKCF-LLEQLQVISHESGDTDIHD 61

QY 92 LVAQ---LRRRLPARRGKKQKHIAKAPCSDSYSEKRTPKPELERIKWLLQXMIH 142

Db 62 TVENLIILANNILSSNGNIT- --SGKCECELEEKNIKEPQSFVHIQVOMFIN 112
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

```

RESULT 5
US-08-393-305-6
; Sequence 6, Application US/08393305
; Patent No. 5574138
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,305
; FILING DATE: 22-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.

```

```
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-393-305-6

Query Match      8.9%; Score 68.5; DB 1; Length 114;
Best Local Similarity 22.8%; Pred. No. 0.63;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps

Qy    32 LRLHLDIVLEQLKIYENDLDPDLLSAPODVKGHCHEAAFCOKAKLKPSNPNKKTFTID 91
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     9 LKKIEDLIQSMMI-----DATLYTESDVHFSCVKTAMKCF-LLELVQVISHESGDTDHD 61

Qy    92 LVQAQ---LRRKLPRARGKKKHIAKACSPCDSYEKRTPKPEFLERLUWLQKMIH 142
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db    62 TVENLLIIANNILSSNGHITE---SGCKCECELEBKNIKEQLQSFVHVQMFIN 112
```

```

RESULT 6
US-08-726-817-6
; Sequence 6, Application US/08726817
; Patent No. 5707616
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eismann, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,817
; FILING DATE: 04-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,305
; FILING DATE: 22-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-726-817-6

```

Query Match 8.9%; Score 68.5; DB 1; Length 114;

RESULT 8

STREET: 6300 Columbia Center, 701 Fifth Avenue

STREET: 6300 Columbia Center, 701 Fifth Avenue

```

; APPLICATION NUMBER: US 08/393,305
; FILING DATE: 22-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McMasters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 480052.409C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-189-193-6

Query Match      8.9%; Score 68.5; DB 4; Length 114;
Best Local Similarity 22.8%; Pred. No. 0.63;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps 4;

QY 32 LRHLIDIVEQLKIYENDLDPELLSAPODVKGCHCHAAFACFQKAKLKPSPNCGNKKTFIID 91
   | : | : | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9 LKKTIEDIQSMHI-----DATLYTESDVHPSCVTAMKCF-LLEQLQVISHESGDTDIHD 61
   | : | : | : | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 92 LVAQ---LRRLPARRGGKKQKHIAKPCSDSYEKRTPKFEFLERLKLWLLQKMIH 142
   | : | : | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 62 TVENLIILANILSSNGNITP---SGKCEEELEKNIKEFLQSFVHIVQMFIN 112
   | : | : | : | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
PCT-US94-03793-3
; Sequence 3, Application PC/TUS9403793
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Interleukin-15
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/03793
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Launer, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
PCT-US94-03793-3

```

```
Query Match      8.9%; Score 68.5; DB 5; Length 114;
Best Local Similarity 22.8%; Pred. No. 0.63;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps 4;

QY 32 LRHLIDIVEQLKIYENDLPDPELLSAPQDVKGHCHEAFAFQKAKLPSPNGNKKTFIID 91
Db 9 LKKIEDLIQSMHI-----DATLYTESDVHPSCVKVTAMKCF-LLELQVISHESGDTDIHD 61
QY 92 LVAQ---LRRRLPARRGGKKOKHIAKPCSDSYEKRTPKFEFLERLKWLLQKMIH 142
Db 62 TVENLIILANLSSNGNITE---SGCKECELEEKNIKEFLQSFVHVQMFIN 112

RESULT 12
US-08-300-903A-3
; Sequence 3, Application US/08300903A
; Patent No. 5591630
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M
; APPLICANT: Girl, Judith G
; TITLE OF INVENTION: Interleukin-15 Receptors
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple Operating System 7.1
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/300,903A
; FILING DATE: 06-SEPTEMBER-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/236,919
; FILING DATE: 06-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2822-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; TELEFAX: 206-233-0644
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-300-903A-3

Query Match      8.9%; Score 68.5; DB 1; Length 122;
Best Local Similarity 22.8%; Pred. No. 0.69;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps 4;

QY 32 LRHLIDIVEQLKIYENDLPDPELLSAPQDVKGHCHEAFAFQKAKLPSPNGNKKTFIID 91
Db 17 LKKIEDLIQSMHI-----DATLYTESDVHPSCVKVTAMKCF-LLELQVISHESGDTDIHD 69
QY 92 LVAQ---LRRRLPARRGGKKOKHIAKPCSDSYEKRTPKFEFLERLKWLLQKMIH 142
Db 70 TVENLIILANLSSNGNITE---SGCKECELEEKNIKEFLQSFVHVQMFIN 120

RESULT 13
```

```
US-08-031-399-2
; Sequence 2, Application US/08031399
; Patent No. 5552303
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: Epithelium-derived T-cell Factor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/031,399
; FILING DATE: 19930308
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauener, Charlene
; REGISTRATION NUMBER: 33,035
; REFERENCE/DOCKET NUMBER: 2811
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-031-399-2

Query Match      8.9%; Score 68.5; DB 1; Length 162;
Best Local Similarity 22.8%; Pred. No. 1;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps 4;

QY 32 LRHLIDIVEQLKIYENDLPDPELLSAPQDVKGHCHEAFAFQKAKLPSPNGNKKTFIID 91
Db 57 LKKIEDLIQSMHI-----DATLYTESDVHPSCVKVTAMKCF-LLELQVISHESGDTDIHD 109
QY 92 LVAQ---LRRRLPARRGGKKOKHIAKPCSDSYEKRTPKFEFLERLKWLLQKMIH 142
Db 110 TVENLIILANLSSNGNITE---SGCKECELEEKNIKEFLQSFVHVQMFIN 160

RESULT 14
US-08-393-305-5
; Sequence 5, Application US/08393305
; Patent No. 5574138
; GENERAL INFORMATION:
; APPLICANT: Grabstein, Kenneth
; APPLICANT: Anderson, Dirk
; APPLICANT: Eisenman, June
; APPLICANT: Fung, Victor
; APPLICANT: Rauch, Charles
; TITLE OF INVENTION: EPITHELIUM-DERIVED T-CELL FACTOR
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
```

```

, ZIP: 98104
,
, COMPUTER READABLE FORM:
,
, MEDIUM TYPE: Floppy disk
,
, COMPUTER: IBM PC compatible
,
, OPERATING SYSTEM: PC-DOS/MS-DOS
,
, SOFTWARE: PatentIn Release #1.0, Version #1.25
,
, CURRENT APPLICATION DATA:
,
, APPLICATION NUMBER: US/08/393,305
,
, FILING DATE: 22-FEB-1995
,
, CLASSIFICATION: 424
,
, ATTORNEY/AGENT INFORMATION:
,
, NAME: McMasters, David D.
,
, REGISTRATION NUMBER: 33,963
,
, REFERENCE/DOCKET NUMBER: 480052.409C2
,
, TELECOMMUNICATION INFORMATION:
,
, TELEPHONE: 206-622-4900
,
, INFORMATION FOR SEQ ID NO: 5:
,
, SEQUENCE CHARACTERISTICS:
,
, LENGTH: 162 amino acids
,
, TYPE: amino acid
,
, TOPOLOGY: linear
,
, MOLECULE TYPE: protein
,
, US-08-393-305-5

```

	Query Match	8.9%; Score 68.5; DB 1; Length 162;
	Best Local Similarity	22.8%; Pred. No. 1;
	Matches	26; Conservative 22; Mismatches 53; Indels 13; Gaps 4;
QY	32 LRHLIDIVEQLKIYENDLDPBELLISAPODVVGHCHEAFACFQAKLPSNPGNNKTFIID	91
	: : : : : : : :	
Db	57 LKKIEDLIQSMHI-----DATLYTESDVHPSCVKVTAMKCF-LLEQLQVISHESGDTIHD	109
QY	92 LVAQ---LRRLPARRGGKKQHIAKCPCSDCSYEKRTPKFEFLERKLWLLQXMIH	142
	:	
Db	110 TVENLIILANILLSSNGITE----SGCKEEELKEKNIKFELQSFVHVQFMFIN	160

Search completed: May 23, 2001, 11:44:17
Job time: 1254 sec

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-284-393B-9

Query Match      8.9%; Score 68.5; DB 1; Length 162;
Best Local Similarity 22.8%; Pred. No. 1;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps

Qy 32 LRHLIDIVEQLKIYENDLDPellisAQDVKGCEHAAPCFQAKLKLPSPGPKTKFTIID 91
   | : : : : : | | | | | | | | | | | | | | | | | | | | | |
Db 57 LKIIDLIQSMHI-----DATLYTESDVHPSCKVTAMKCF-LLEQVISHESGDDIID 109
   | : : : : : | | | | | | | | | | | | | | | | | | | | | |

Qy 92 LVAQ----LRRLPARRGGKKQKHAICPCSDSYEKRTPEFLERLKLWLQKMIH 142
   | : : : : : | | | | | | | | | | | | | | | | | | | | | |
Db 110 TVENLIILANNILSSNGNITE---SGCKEECELEEKNIKEFLOSFVHIOMFIN 160
   | : : : : : | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 15
US-08-284-393B-9
; Sequence 9, Application US/08284393B
; Patent No. 596234
; GENERAL INFORMATION:
; APPLICANT: Zurawski, Sandra M.
; APPLICANT: Zurawski, Gerard
; TITLE OF INVENTION: MUTINS OF MAMMALIAN CYTOKINES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,393B
; FILING DATE: 01-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0389
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-852-9196
; TELEFAX: 415-496-1200
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 amino acids

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-284-393B-9

Query Match      8.9%; Score 68.5; DB 1; Length 162;
Best Local Similarity 22.8%; Pred. No. 1;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps

Qy 32 LRHLIDIVEQLKIYENDLDPellisAQDVKGCEHAFACFQAKLKPSNPGNKKFTIID 91
   | : : : : : | | | | | | | | | | | | | | | | | | | | | |
Db 57 LKIIDLIQSMHI-----DATLYTESDVHPSCKVTAMKCF-LLEQVISHESGDDIID 109

Qy 92 LVAQ----LRRLPARRGGKKQKHAICPCSDSYEKRTPEFLERLKLWLQKMIH 142
   | : : : : | | | | | | | | | | | | | | | | | | | | | |
Db 110 TVENLIILANNILSSNGNITE---SGCKEECELEEKNIKEFLOSFVHI VOMFIN 160
   | : : : : | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: May 23, 2001, 11:44:17
Job time: 1254 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:23:23 ; Search time 57.43 Seconds
(without alignments)
174.709 Million cell updates/sec

Title: US-09-522-217-56
Perfect score: 769
Sequence: 1 MERTVLVCLVIFLGTVAHKS.....KEFLERLKWLLQKMIHQHLS 146

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_67:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	82.5	10.7	414	2	S55654	DNA polymerase pro
2	78.5	10.2	404	2	T27106	hypothetical prote
3	78.5	10.2	440	2	T47906	FUSCA PROTEIN FUS6
4	75	9.8	718	1	A69084	DNA topoisomerase
5	74	9.6	1081	2	T18223	histidine kinase -
6	74	9.6	1081	2	T18258	histidine kinase -
7	74	9.6	1081	2	T14003	histidine kinase h
8	74	9.6	1081	2	T14173	histidine kinase h
9	74	9.6	1142	2	A45031	cysteine-rich fibr
10	73	9.5	156	2	T10140	modA.3 protein - p
11	72.5	9.4	510	2	JS0724	cytochrome P450 AL
12	72.5	9.4	698	2	S25409	transcription fact
13	71.5	9.3	403	2	B69196	conserved hypothet
14	71	9.2	409	2	E83992	ATP/GTP-binding pr
15	70.5	9.2	253	2	T40336	heterodisulfide re
16	70.5	9.2	324	2	S54361	gag-related protei
17	70.5	9.2	487	2	JC2495	histamine H1 recep
18	70.5	9.2	874	2	A64664	valine--tRNA ligas
19	70.5	9.2	1127	1	GNWVD2	genome polyprotein
20	69.5	9.0	1324	2	S51622	cut3 protein - fls
21	69	9.0	450	2	S00950	hypothetical prote
22	69	9.0	855	2	A34810	3',5'-cyclic-GMP p
23	68.5	8.9	162	1	A53484	interleukin-15 pr
24	68.5	8.9	293	2	T31146	probable DNA inver
25	68.5	8.9	472	2	T19700	hypothetical prote
26	68.5	8.9	491	2	A41632	histamine H1 recep
27	68	8.8	353	2	T15429	hypothetical prote
28	68	8.8	1256	2	C71436	probable resistanc
29	68	8.8	1603	2	S23810	collagen alpha 1(X

30 68 8.8 2749 1 ACMSIT inositol 1,4,5-tri
31 67.5 8.7 1205 2 T41987 hypothetical prote
32 67 8.7 379 2 F82450 probable exonuclea
33 66.5 8.6 355 2 JC1249 peroxidase (EC 1.1
34 66.5 8.6 364 2 S34355 peroxidase (EC 1.1
35 66.5 8.6 443 2 D82975 two-component sens
36 66.5 8.6 829 2 T26969 hypothetical prote
37 66.5 8.6 853 2 H70470 hypothetical prote
38 66 8.6 128 2 T32947 hypothetical prote
39 66 8.6 448 2 JQ1161 Gag protein - Maed
40 66 8.6 452 2 D64961 membrane protein y
41 66 8.6 617 2 JC5721 vacuolar protein s
42 66 8.6 789 2 PC4161 zinc finger protei
43 66 8.6 1045 2 S60571 integrin alpha v c
44 65.5 8.5 329 2 T21038 hypothetical prote
45 65.5 8.5 377 2 T16764 hypothetical prote

ALIGNMENTS

RESULT 1
S55654
DNA polymerase processivity chain - equine herpesvirus 2
C:Species: equine herpesvirus 2
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 08-Oct-1999
C:Accession: S55654
R:Telford, E.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.
J. Mol. Biol. 249, 520-528, 1995
A:Title: The DNA sequence of equine herpesvirus 2.
A:Reference number: S55594; MUID:95302501
A:Accession: S55654
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-414 <TEL>
A:Cross-references: GB:U20824; NID:g995172; PIDN:AAC13847.1; PID:g695232
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995
C:Superfamily: human herpesvirus 4 BHRF1 protein

Query Match 10.7%; Score 82.5; DB 2; Length 414;
Best Local Similarity 25.5%; Pred. NO. 2.1;
Matches 28; Conservative 16; Mismatches 51; Indels 15; Gaps 3;

QY 19 KSSPQGPDRLLRLRLHLDIVLEQLKIYENDLDPPELLSAPQDVKGHCHEAAFACFQKAKLK 78
DB 281 KSSKPASGELSFLLRANPQVDFNGVQEQVSS---VASTCRHLSESCSLDPPPT 337
QY 79 PSNPGNNKTFIIDLVAQLRRRLPARRGGKKQKHIAKCPSCDSYERKTPKE 128
DB 338 PELPGSPDTF-----KEIPGRSG---SVHLERDLSCSDSEETPKQ 375

RESULT 2
T27106
hypothetical protein Y52B11A.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27106
R:Lennard, N.
A:Reference number: Z20310
A:Accession: T27106
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-404 <WTL>
A:Cross-references: EMBL:AL032654; PIDN:CAA21720.1; GSPDB:GN00019; CESP:Y52B11A.9
C:Genetics:
A:Gene: CESP:Y52B11A.9
A:Map position: 1
A:Introns: 26/2; 102/3; 301/3; 367/3

Query Match 10.2%; Score 78.5; DB 2; Length 404;
 Best Local Similarity 23.8%; Pred. No. 5;
 Matches 34; Conservative 20; Mismatches 58; Indels 31; Gaps 6;

QY 14 GTVAHSSPOGRLRLRLRLHLDIVQLKIYENDLPELLSAPODVKGHCHEAFAACFQ 73
 DB 173 GWQRGELAGDDEHEVEATLRTDPQKI-QLDNLGLDRKLDVKSGVASAKISFD 231

QY 74 KAKL-----PSNPGNNKTFIIDLVAQLRRRLPARRGGKQKHIACPCSDSYEKRTF----- 126
 DB 232 MPVKKEDPDPEGPSQ-----PSRSGKKRSR-SRSPAARFKFSKKSALDEI 276

QY 127 KEFLERLK-----WLLQKMI 141
 DB 277 KEMERKKRKNRKYWMREGIV 299

RESULT 3
 T47906
 FUSCA PROTEIN FUS6 - Arabidopsis thaliana
 N:Alternate names: protein T20K12.40
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T47906
 R:De Haan, M.; Maarse, A.C.; Grivell, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quet
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: 224480
 A:Accession: T47906
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-440 <DEH>
 A:Cross-references: EMBL:AL137898
 A:Experimental source: cultivar Columbia; BAC clone T20K12
 C:Genetics:
 A:Map position: 3
 A:Introns: 130/3; 237/3; 278/1; 304/2; 382/3
 A:Note: T20K12.40

Query Match 10.2%; Score 78.5; DB 2; Length 440;
 Best Local Similarity 25.9%; Pred. No. 5.4; Indels 51; Gaps 7;

QY 8 LVVIFLGTVAHKSS-----PQGPD-----RLRLRLHLDIVQLKIYE----- 46
 DB 184 LVSTEMGQFHTVTSYVKNKAEQNPETLEPMVNAKLRCSGLAHL-----ELKYYKLAARKF 238

QY 47 NLDPEL-----LSAPQDVKGHCHEAFAACFQKAKLPSNPGNNKTFIIDLVAQLRRRL 100
 DB 239 LDVNPGLNGSYNEVIAQDIATYGGICALASFDSELKAFIDNFRNLFELVDPVRELI 298

QY 101 PARGGKKQKHIACPCSDSYEKRTF--EFLERLKWLLQKMIHQH 144
 DB 299 N-----DFYSRYASCLEYLAKSLNLLDILHL 327

RESULT 4
 A69084
 DNA topoisomerase I - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A69084
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.
 ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
 A:Reference number: A69000; MUID:98037514
 A:Accession: A69084
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-718 <MTH>

A:Cross-references: GB:AE000921; GB:AE000666; NID:g2622744; PIDN:AAB86097.1; PID:g262
 A:Experimental source: Strain Delta H
 C:Genetics:
 A:Gene: MTH1624
 A:Start codon: GTG
 C:Superfamily: DNA topoisomerase I topa

Query Match 9.8%; Score 75; DB 1; Length 718;
 Best Local Similarity 26.1%; Pred. No. 19;
 Matches 35; Conservative 16; Mismatches 35; Indels 48; Gaps 8;

QY 26 DRLRLRLHLDIVQLKIYENDLPELLSAPODVKGHCHEAFAACFQKAKLPSNP 82
 DB 561 DEATEVRSILSDIENLR-----DIGKELYRAYQDSRVVGECP-----ACGGKLVIKIYS-P 611

QY 83 GNNKTFI-----IDLVAQLRRR-----LPARRGGKK- 108
 DB 612 RNRSTFVGCSSYPDCRTVYSLPRGASVLKSLCEKGLPMISYGRPRQACLDPKCKGKKKS 671

QY 109 --QKHIACPCSDS 120
 DB 672 EVEEVGKCEGCS 685

RESULT 5
 T18223
 histidine kinase - yeast (Candida albicans)
 C:Species: Candida albicans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T18223
 R:Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, November 1998
 A:Reference number: Z18831
 A:Accession: T18223
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1081 <BAR>
 A:Cross-references: EMBL:AL033396; NID:e1340089; PID:e1340096; PIDN:CAA21950.1
 C:Genetics:
 A:Note: Ca35A5.07

Query Match 9.6%; Score 74; DB 2; Length 1081;
 Best Local Similarity 27.6%; Pred. No. 36;
 Matches 34; Conservative 19; Mismatches 50; Indels 20; Gaps 7;

QY 27 RLILRLHLDIVQLKIYE-----NLDPELLSAPODVKGHCHEAFAACFQKAKL--KP 79
 DB 891 RILEKGHSVEVENGLEAYEAIKRNKYDVLMQVPMYMG-----GFEATEKIQWEKK 945

QY 80 SNPGNNKTFIIDLVAQLRRRLPARRGGKKQKHIACPCSDSYEKRTFKEFLERLKWLLQK 139
 DB 946 SNPDSLTFRTPIIALTAHAML-----GDREKSLAK--GMDDYVSKPLKPKL--LMOTINK 997

QY 140 MTH 142
 DB 998 CIH 1000

RESULT 6
 T18258
 histidine kinase - yeast (Candida albicans)
 C:Species: Candida albicans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T18258
 R:Alex, L.A.; Korch, C.; Selitrennikoff, C.P.; Simon, M.I.
 Proc. Natl. Acad. Sci. U.S.A. 95, 7069-7073, 1998
 A:Title: COS1, a two-component histidine kinase that is involved in hyphal developmen
 A:Reference number: Z18845; MUID:98284059
 A:Accession: T18258
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

C; Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 01-Dec-2000

C; Accession: S25409

R; Lloyd, S.L.; Sargent, C.A.; Chalmers, J.; Lim, E.; Habeebu, S.S.M.; Affara, N.A.

Nucleic Acids Res. 19, 4835-4841, 1991

A; Title: An X-linked zinc finger gene mapping to Xq21.1-q21.3 closely related to ZFX

A; Reference number: S25409; MUID:92020112

A; Accession: S25409

A; Status: preliminary

A; Molecule type: mRNA

A; Residues: 1-698 <LLO>

A; Cross-references: EMBL:X56465; NID:g38027; PIDN:CAA39837.2; PID:g5304925

A; Note: the authors translated the codon CAG for residue 4 as Glu, GAA for residue 53 as Arg, AAG for residue 414 as His, AGT for residue 575 as Val, GAT for residue 576 as A

C; Genetics:

A; Gene: znf6

C; Superfamily: zinc finger protein ZFP-36; LIM metal-binding repeat homology

Query Match 9.4%; Score 72.5; DB 2; Length 698;

Best Local Similarity 26.8%; Pred. No. 32;

Matches 26; Conservative 12; Mismatches 32; Indels 27; Gaps 5;

QY 57 PDQVKGH-----CEHAFACFQKAKLK--PSNPGNNKFTIIDLVAQ----- 95

DB 454 PSELKKMRHTGEGKPYQCYICFRCADQNLKTHIKSGNNLPYKCEHCPOAFGDRE 513

QY 96 LRRRLPARRGGKKOKHIAKPCSDSYEKRTPKPEFLER 132

DB 514 LQRHLDFQGHKTH---QCPHCD--HKSTNSSDLKR 544

RESULT 13

B69196

C; Species: Methanobacterium thermoautotrophicum (strain D)

C; Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 11-Jan-2000

C; Accession: B69196

R; Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, J. S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A; Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu

A; Reference number: A69000; MUID:98037514

A; Accession: B69196

A; Status: preliminary; nucleic acid sequence not shown; translation not shown

A; Molecule type: DNA

A; Residues: 1-403 <MTH>

A; Cross-references: GB:AE000798; GB:AE000666; NID:g2621094; PIDN:AAB84576.1; PID:g2621094

A; Experimental source: strain Delta H

C; Genetics:

A; Gene: MTH72

C; Superfamily: unassigned tetratricopeptide repeat proteins; tetratricopeptide repeat

F; 16-49/Domain: tetratricopeptide repeat homology <TT01>

F; 50-83/Domain: tetratricopeptide repeat homology <TT02>

F; 84-117/Domain: tetratricopeptide repeat homology <TT03>

F; 118-151/Domain: tetratricopeptide repeat homology <TT04>

F; 152-185/Domain: tetratricopeptide repeat homology <TT05>

F; 186-219/Domain: tetratricopeptide repeat homology <TT06>

F; 220-253/Domain: tetratricopeptide repeat homology <TT07>

F; 254-287/Domain: tetratricopeptide repeat homology <TT08>

F; 288-321/Domain: tetratricopeptide repeat homology <TT09>

F; 322-355/Domain: tetratricopeptide repeat homology <TT10>

F; 356-389/Domain: tetratricopeptide repeat homology <TT11>

Query Match 9.3%; Score 71.5; DB 2; Length 403;

Best Local Similarity 27.0%; Pred. No. 23;

Matches 33; Conservative 22; Mismatches 32; Indels 35; Gaps 8;

QY 17 AHKSSGPGORLLRLRLHLIDI---VSQLIYENDL--DPSELLSA-----PDQVKGHC 64

DB 42 ALKASPNDDTLYHNTMLTKLKRPEALKCYEKLKNNPKLAENWNNKGVWLKELKRYD 101

Qy	37	DIVEOLKVIENDLD	---	PELLSAPODVKGCHBEAFACF	---	---	ORAKLPSNPGNNKTF	88
Db	52	DVISDLKIIDKIIQSLHIDATL	TESDVHPNCKVTAMKCF	LLHVLHVISLESKNETHTQV	111			
Qy	89	--IIDLVQAOLRRRLPARRG	GKKQKHIAK	--CPSCDSYEKETPEFETLERLKLWLQKMIH	142			
Db	112	ENIIIL	-----	ANSGLSSNRNITETCSCKE	EELEKNEIKFELFVHVHIVOMFIN	160		

```

RESULT 2
FUS6_ARATH STANDARD; PRT; 441 AA.
AC P45432;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FUSCA PROTEIN FUS6.
GN FUS6 OR COP11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. WASSILEWSKIJA; TISSUE=Siliques;
RX MEDLINE=94176998; PubMed=8130643;
RA Castle L.A., Meinke D.W.;
RT "A FUSCA gene of Arabidopsis encodes a novel protein essential for
  plant development."
RL Plant Cell 6:25-41(1994).
CC -!- FUNCTION: IS AN ESSENTIAL REGULATORY PROTEIN. HAS AN ESSENTIAL
  ROLE THROUGHOUT PLANT DEVELOPMENT. IT COULD BE IMPORTANT IN
  RELAYING PLANT HORMONE AND OTHER SIGNALS FROM MEMBRANE-ASSOCIATED
  PKC TO CYTOPLASMIC FACTORS AND THIS COULD LEAD DIRECTLY TO
  BIOCHEMICAL RESPONSES OR TO DOWNSTREAM GENE REGULATION.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC -!- TISSUE SPECIFICITY: LEAVES, FLOWERS, IMMATURE SILIQUES, AND
  LIGHT-GROWN ROOTS.
CC -!- SIMILARITY: BELONGS TO THE FUS6 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L26498; AAA32792.1; -
DR InterPro; IPR000717; -
DR Pfam; PF01399; PCI; 1.
SQ SEQUENCE 441 AA; 50548 MW; 2F6F223FCC337756 CRC64;

Query Match 9.9%; Score 76; DB 1; Length 441;
Best Local Similarity 26.3%; Pred. No. 2.7;
Matches 44; Conservative 18; Mismatches 53; Indels 52; Gaps 8;

QY 8 LVVIFLGTVAHKSS-----PQGPD-----RLIRLRHLIDIVEQLKIYE----- 46
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 184 LVSIEMGFQTHVTYVYVKAQNPETLEPMVNAKLRCASGLAHL-----ELKKYLAARKF 238

QY 47 NDLPDEL-----LSAPQDVGKHCHEAAFCQAKLKPSNPGN-NKTFIIDIVLAQLRRR 99
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 239 LDVNPGLNSYNEVIAPIQATYGGGLCALASFDRELKQKVIDINFRNFLELVPDREL 298

QY 100 LPARRGKKQKHIAKPCSDSYKRTPK--EFLERLKWLLQKMIHQH 144
  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||  :|||
DB 299 IN-----DFYSRYASCLEYLASLNLLLDIHLH 328

RESULT 3
Y01F_BPT4
ID Y01F_BPT4 STANDARD; PRT; 156 AA.
AC P39425;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 18.3 KDA PROTEIN IN MODB-MRH INTERGENIC REGION.
  .

Y01F OR DDA.7.
OS Bacteriophage T4.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae;
OC T4-like phages.
OX NCBI_TaxID=10665;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90255970; PubMed=1692800;
RA Frazier M.W., Mosig G.;
RT "The bacteriophage T4 gene mrh whose product inhibits late T4 gene
  expression in an Escherichia coli rpoH (sigma 32) mutant."
RL Gene 88:7-14(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M30001; AAB07799.1; -
DR Hypothetical protein.
KW SEQUENCE 156 AA; 18331 MW; 7AB2F4957156D089 CRC64;

Query Match 9.5%; Score 73; DB 1; Length 156;
Best Local Similarity 29.6%; Pred. No. 1.6;
Matches 21; Conservative 15; Mismatches 25; Indels 10; Gaps 3;

QY 25 PDLRLRLRLHLDIVLEQLKIYENDLPDLSAPQVKGH---CEHAFAACFQAKLK-PS 80
  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 50 PENIILIKHTEDIL-----QNTDSPSSSEALTIKGYKRAHEYGLDFDMDKVKLAS 103

QY 81 NPGNNKTFIID 91
  | :|||:
DB 104 QPSKSKTFIIE 114

RESULT 4
CP5Q_CANNA STANDARD; PRT; 510 AA.
AC Q12587;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CYTOCHROME P450 52C2 (EC 1.14.14.-) (CYPLIIC2) (ALKANE-INDUCIBLE
  P450-ALK6-A).
DN CYP52C2.
OS Candida maltosa (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5479;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95163275; PubMed=7865134;
RA Okuma M., Muraoka S., Tanimoto T., Fujii M., Ohta A., Takagi M.;
RT "CYP52 (cytochrome P450alk) multigene family in Candida maltosa:
  identification and characterization of eight members."
RL DNA Cell Biol. 14:163-173(1995).
CC -!- FUNCTION: TOGETHER WITH AN NADPH CYTOCHROME P450 THE ENZYME SYSTEM
  CATALYZES THE TERMINAL HYDROXYLATION AS THE FIRST STEP IN THE
  ASSIMILATION OF ALKANES AND FATTY ACIDS.
CC -!- CATALYTIC ACTIVITY: HYDROXYLATION OF N-ALKANES AT THE TERMINAL
  POSITION.
CC -!- INDUCTION: BY VARIOUS ALKANES.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

CC EMBL; D12718; BAA02212.1; -
 CC InterPro; IPR001128; -
 CC InterPro; IPR002974; -
 CC Pfam; PF00067; P450; 1; IYCYPS2.
 CC PRINTS; PR01239; EP450IICYPS2.
 CC PROSITE; PS00086; CYTOCHROME_P450; 1.
 CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme.
 FT BINDING 458 458 HEME (BY SIMILARITY).
 SQ SEQUENCE 510 AA; 58113 MW; 2AD3F6B48AC596B8 CRC64;

Query Match 9.4%; Score 72.5; DB 1; Length 510;
 Best Local Similarity 23.1%; Pred. No. 7.3;
 Matches 28; Conservative 20; Mismatches 52; Indels 21; Gaps 4;

QY 28 LLRLRLHLDIVQLKIYENDLPDLSAPQDVKGHCHEHAFAFQAKLKPSNPGNNKT 87

Db 325 LFFELSHENVTTLKEVDPQSPDVESITFTIQNDYLRWCLFESLRVNPSPVFNST 384

QY 88 FIDLVLAQLRRRLPARGGKKQKHIACPCSD-----SYEKRTPKFEFLERLK 134

Db 385 ANKDTI-----LP-RGGGDCSHPLVKGQDVLFPLFYASNRQKRYFGRKPEFIPE-R 436

QY 135 W 135

Db 437 W 437

RESULT 5

H2AY_HUMAN

ID H2AY_HUMAN STANDARD; PRT; 371 AA.

AC 075367; 075377; Q9UP96;

DT 01-OCT-2000 (Rel. 40, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE CORE HISTONE MACRO-H2A.1 (HISTONE MACROH2A1) (MH2A1) (H2A.Y) (H2A.Y).

GN H2AFY OR MACROH2A1

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

[1]

RN SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

RC TISSUE=Liver;

RX MEDLINE=98390273; PubMed=9714746;

RA Lee Y., Hong M., Kim J.W., Hong Y.M., Choe Y.-K., Chang S.Y.,

RA Lee K.S., Choe I.S.;

RT "Isolation of cDNA clones encoding human histone macroH2A1 subtypes.";

RL Biochim. Biophys. Acta 1399:73-77(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Blood;

RX MEDLINE=98318631; PubMed=9653160;

RA Mao M., Fu G., Wu J.-S., Zhang Q.-H., Zhou J., Kan L.-X., Huang Q.-H.,

RA He K.-L., Gu B.-W., Han Z.-G., Shen Y., Gu J., Yu Y.-P., Xu S.-H.,

RA Wang Y.-X., Chen S.-J., Chen Z.;

RT "Identification of genes expressed in human CD34(+) hematopoietic

stem/progenitor cells by expressed sequence tags and efficient full-

length cDNA cloning.";

RL Proc. Natl. Acad. Sci. U.S.A. 95:8175-8180(1998).

CC -!- FUNCTION: MACRO-H2A OCCUPIES THE PLACE OF CONVENTIONAL H2A

IN A SUBSET OF NUCLEOSOMES AND WHILE IT HAS RETAINED GENERAL

H2A FUNCTION, IT MAY ALSO HELP TO PRODUCE SEQUENCE-SPECIFIC

NUCLEOSOME POSITIONING.

CC -!- SUBCELLULAR LOCATION: NUCLEAR.

CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 AND 2 (SHOWN HERE); ARE

PRODUCED BY ALTERNATIVE SPLICING.

CC -!- TISSUE SPECIFICITY: UBUIQUITOUS.

CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE HISTONE H2A

FAMILY.

CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE DUF27

FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

DR EMBL; AF041483; AAC33433.1; -

DR EMBL; AF044286; AAC33434.1; -

DR EMBL; AF054174; AAC39908.1; -

DR InterPro; IPR000166; -

DR InterPro; IPR002119; -

DR InterPro; IPR002589; -

DR Pfam; PF00125; histone; 1.

DR Pfam; PF01661; DUF27; 1.

DR PRINTS; PR00620; HISTONEH2A.

KW Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding;

Multigene family; Alternative splicing.

FT INIT_MET 0 0 BY SIMILARITY.

FT DOMAIN 1 116 HISTONE H2A.

FT DOMAIN 117 161 LYS-RICH.

FT DOMAIN 215 330 DUF27.

FT BINDING 116 116 UBUIQUITIN (BY SIMILARITY).

FT VARSPIC 197 228 NLIHSEISNLGAFVEALINPTNADIDPKDDL -> OVVOA

FT CONFLICT 160 160 MISSING (IN REF. 2).

FT CONFLICT 224 224 P -> L (IN REF. 2).

SQ SEQUENCE 371 AA; 39469 MW; 3EFAD151C2500535 CRC64;

Query Match 9.2%; Score 70.5; DB 1; Length 371;

Best Local Similarity 29.5%; Pred. No. 8;

Matches 31; Conservative 16; Mismatches 45; Indels 13; Gaps 5;

QY 48 DLDPELLSAPQDVKGHCHEHAFAFQAKLKPSNPGNNKTFFIDLVLAQLRRRLPARRGCK 107

Db 107 NIHELLAKKSGSKGLE--AIITPPAK-KAKSPQKK-----PVSKKAGGKKGAKSKK 159

QY 108 KQKIHAQPCSDSYEKRTPKFEFLERLK---WLLQK--MIHQHLS 146

Db 160 KQGEVSKAASADSTEGTPADGFTVLSTKSLFLGKLNLIHSEIS 204

RESULT 6

H1R_HUMAN

ID H1R_HUMAN STANDARD; PRT; 487 AA.

AC P35367;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE HISTAMINE H1 RECEPTOR.

GN HRH1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94107375; PubMed=8280179;

RA de Backer M.D., Gommeren W., Moereels H., Nobels G., van Gompel P.,

RA Leysen J.E., Luyten W.H.M.L.;

RT "Genomic cloning, heterologous expression and pharmacological

characterization of a human histamine H1 receptor.";

RL Biochem. Biophys. Res. Commun. 197:1601-1608(1993).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=94271250; PubMed=8003029;

RA Fukui K., Fujimoto K., Mizuguchi H., Sakamoto K., Horio Y., Takai S.,

RA Yamada K., Ito S.;

Db 756 VGLREAENTERLQTYAOKLARLEKYSVSSKPLKSVSDVGEFCQ--TYANLENLDLSP- 812

QY 81 NPGNNKTFIIDLVAQRRRLPARGGKKQKHIAKPCSDSYEKRTPKFEFLER-----LKWL 136

Db 813 -----LVARLKKQLEKEKELKLNHN-----ENFVNAPKSVLEKAKESLKL 857

QY 137 LQK 139

Db 858 LEK 860

RESULT 8

POLG_DEN2D STANDARD; PRT; 1127 AA.

ID POLG_DEN2D 30026; Q66450;

AC P30026; Q66450;

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE GENOME POLYPROTEIN [CONTAINS: CAPSID PROTEIN C (CORE PROTEIN); MATRIX PROTEIN (ENVELOPE GLYCOPROTEIN M); MAJOR ENVELOPE PROTEIN E; NONSTRUCTURAL PROTEIN NS1] (FRAGMENT).

DE Denque virus type 2 (strain D2-04).

OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Flavivirus.

OX NCBI_TaxID=31636;

RN [1]

RP SEQUENCE OF 1-775 FROM N.A.

RA Yang P.Y., Lam S.K.;

RT "The nucleotide and encoded amino acid sequences of the structural protein gene of D2-04 virus strain isolated in China.";

RL Chin. J. Microbiol. Immunol. 11:341-344(1991).

RN [2]

RP SEQUENCE OF 776-1127 FROM N.A.

RA Yang P.Y., Kautner I.M., Koh C.L., Lam S.K.;

RT "Nucleotide and encoded amino acid sequences of the nonstructural protein NS1 gene of a Dengue-2 virus isolated in China.";

RL Chin. J. Microbiol. Immunol. 11:9-12(1991).

CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; U19778; AA73471.1; -

DR PIR; JC1007; GNWD2.

DR HSSP; P14336; LSVB.

DR InterPro; IPR000069; -

DR InterPro; IPR000336; -

DR InterPro; IPR001122; -

DR InterPro; IPR001157; -

DR InterPro; IPR002535; -

DR Pfam; PF01004; Flavi_M; 1.

DR Pfam; PF00948; Flavi_NS1; 1.

DR Pfam; PF01003; Flavi_capsid; 1.

DR Pfam; PF00869; Flavi_glycoprot; 1.

DR Pfam; PF01570; Flavi_propep; 1.

KW Polyprotein; Glycoprotein; Core protein; Coat protein; Envelope protein; Transmembrane; Nonstructural protein.

FT CHAIN 1 114 CAPSID PROTEIN C.

FT PROPEP 115 205 ENVELOPE GLYCOPROTEIN M.

FT CHAIN 206 280 ENVELOPE PROTEIN E.

FT CHAIN 281 775 NONSTRUCTURAL PROTEIN NS1.

FT CHAIN 776 1127 POTENTIAL.

FT TRANSMEM 101 117 POTENTIAL.

FT TRANSMEM 727 743 POTENTIAL.

FT TRANSMEM 757 773 POTENTIAL.

FT DISULFID 283 310 BY SIMILARITY.

FT DISULFID 340 396 BY SIMILARITY.

FT DISULFID 354 385 BY SIMILARITY.

FT DISULFID 372 401 BY SIMILARITY.

FT DISULFID 465 565 BY SIMILARITY.

FT DISULFID 582 613 BY SIMILARITY.

FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 905 905 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 982 982 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT NON_TER 1127 1127

SQ SEQUENCE 1127 AA; 125835 MM; C93D541432DA80EE CRC64;

Query Match 9.2%; Score 70.5; DB 1; Length 1127;

Best Local Similarity 22.0%; Pred No. 29;

Matches 24; Conservative 23; Mismatches 29; Indels 33; Gaps 4;

QY 15 TVAHKSPQGGPDL-----LIRLHLLI--DIVEQLKIYENDLDPILLS 55

Db 804 TEQYNFQPEPSKSLASAMRKAHEEGICGIRSVTRLENLWKQITPELKHLSIEVKLTI 863

QY 56 APQDVKGHCERH-----AFACFQAKLKPSPGNKNKTFIID 91

Db 864 MTGDIKIGIMQAGTRSLRPQPTKELKFSWETWRKAKMVPTEP-HNQTEFLID 911

RESULT 9

CUT3_SCHPO STANDARD; PRT; 1324 AA.

ID CUT3_SCHPO 411004;

AC P411004;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE CHROMOSOME SEGREGATION PROTEIN CUT3.

GN CUT3.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE-95045386; PubMed-7957061;

RA Saka Y., Sutani T., Yamashita Y., Saitoh S., Takeuchi M., Nakaseko Y., Yanagida M.;

RT "Fission yeast cut3 and cut14, members of a ubiquitous protein family, are required for chromosome condensation and segregation in mitosis."

RL EMBO J. 13:4938-4952(1994).

CC -1- FUNCTION: REQUIRED FOR CHROMOSOME CONDENSATION AND SEGREGATION IN MITOSIS.

CC -1- SUBCELLULAR LOCATION: NUCLEAR.

CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; D30788; BAA06454.1; -

KW Mitosis; ATP-binding; Coiled coil; Nuclear protein.

FT NP_BIND 155 162 ATP (POTENTIAL).

FT DOMAIN 310 337 COILED COIL (POTENTIAL).

FT DOMAIN 370 628 COILED COIL (POTENTIAL).

FT DOMAIN 825 1077 COILED COIL (POTENTIAL).

FT DOMAIN 1297 1324 COILED COIL (POTENTIAL).

SQ SEQUENCE 1324 AA; 150594 MM; 304DA5873291F837 CRC64;

```
Query Match          9.0%; Score 69.5; DB 1; Length 1324;
Best Local Similarity 21.5%; Pred. No. 44;
Matches 38; Conservative 19; Mismatches 49; Indels 71; Gaps 6;

QY 23 QGDRLLRLRLHRLDIVEQLKIYENDLDPPELLSAPQDVKGHC-EHAAFACFQAKLKPSN 81
DB 283 EGGDGLLELYEDIIIGTSKYKPIIENM--QELNSDDICAESKRLVLSEKAKLEDK 340
QY 82 PG-----NKKTFIDLVLAQLRRRLPARRGGKKQKHIK 114
DB 341 NSVLSFKDENELFMKQNLRYRLILYETRNKKTIVQLNSVE-----GKLOAHLEK 392
QY 115 CPSCD-----SYEKRTPKFEFLERLKWLLQK 139
DB 393 CEQTERDISKEENEKSLREKAAKVNDCSTSEKTKRQSYEQQTVK-IEEQLKFLNKK 448

RESULT 10
V50K_BYDVP          STANDARD;          PRT;          450 AA.
AC P09516;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DE 01-NOV-1990 (Rel. 16, Last annotation update)
DE 50 KDA PROTEIN (ORF 4).
OS Barley yellow dwarf virus (isolate PAV) (BYDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Luteoviridae;
OC Luteovirus.
OX NCBI_TaxID=12040;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88289355; PubMed=3399386;
RA Miller W.A., Waterhouse P.M., Gerlach W.L.;
RT "Sequence and organization of barley yellow dwarf virus genomic RNA.";
RL Nucleic Acids Res. 16:6097-6111(1988).
CC -1- SIMILARITY: TO ORF 6 OF BEET WESTERN YELLOW VIRUS, ORF6 OF
CC POTATO LEAFROLL VIRUS AND OF 54K PROTEIN OF BEET NECROSIS.
CC YELLOW VEIN VIRUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; X07653; CAA30495.1; -
DR PIR; S00950; S00950.
DR InterPro: IPR002929; -
DR Pfam; PF01690; PLRV_ORF5; 1.
SQ SEQUENCE 450 AA; 49733 MW; 5878FA9361498205 CRC64;

Query Match          9.0%; Score 69; DB 1; Length 450;
Best Local Similarity 27.3%; Pred. No. 14;
Matches 27; Conservative 11; Mismatches 31; Indels 30; Gaps 5;

QY 46 ENLDLP---ELLSAPQDVKGHC-----EHAAFACFQAKLKPSNPNKNTFIIDLVAQLR 97
DB 322 EEDVLPSKEQLSSKPMDSNIIIPKPEPEVLGYQGNITYPE-----DVPPMAR 371
QY 98 RRLPARRGGKKQKHIKCPSCDSYEKTPKE---FLERL 133
DB 372 OKL-----REANAPSTLLYERTPKKSGNFLSRL 401

RESULT 11
CNRC_BOVIN          STANDARD;          PRT;          855 AA.
AC P16586;
DT 01-AUG-1990 (Rel. 15, Created)
```

```
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE CONE CGMP-SPECIFIC 3',5'-CYCLIC PHOSPHODIESTERASE ALPHA'-SUBUNIT
DE (EC 3.1.4.17) (PDE V-C1).
GN PDE6C OR PDEA2
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90115860; PubMed=2153291;
RA Li T., Volpp K., Applebury M.L.;
RT "Bovine cone photoreceptor CGMP phosphodiesterase structure deduced
RT from a cDNA clone.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:293-297(1990).
RN [2]
RP SEQUENCE OF 308-502 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90115859; PubMed=2153290;
RA Charbonneau H., Prusti R.K., Letrong H., Sonnenburg W.K.,
RA Mullaney P.J., Walsh K., Beavo J.A.;
RT "Identification of a noncatalytic CGMP-binding domain conserved in
RT both the CGMP-stimulated and photoreceptor cyclic nucleotide
RT phosphodiesterases.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:288-292(1990).
CC -1- CATALYTIC ACTIVITY: GUANOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O =
CC GUANOSINE 5'-PHOSPHATE.
CC -1- SUBUNIT: COMPOSED OF TWO ALPHA' SUBUNITS THAT ARE ASSOCIATED
CC WITH 3 SMALLER PROTEINS OF 11, 13, AND 15 KDA.
CC -1- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; M37838; AAA30687.1; -
DR EMBL; M33140; AAA30688.1; -
DR PIR; A34810; A34810.
DR InterPro: IPR002073; -
DR InterPro: IPR003018; -
DR Pfam; PF01590; GAF; 2.
DR Pfam; PF00233; PDEase; 1.
DR PRINTS; PR00387; PDIESTERASE1.
DR PROSITE; PS00126; PDEASE_I; 1.
KW Hydrolase; CGMP; Vision; Prenylation; Lipoprotein; Membrane.
FT LIPID 852 852 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 855 AA; 98797 MW; 1FCFFFD045686D65 CRC64;

Query Match          9.0%; Score 69; DB 1; Length 855;
Best Local Similarity 23.1%; Pred. No. 30;
Matches 40; Conservative 20; Mismatches 61; Indels 52; Gaps 6;

QY 13 LGTVAKHSSPOGDRLLRLRLHRLDIV-----EQLKIIY-----ENDL 49
DB 293 LGEVEPYKPTPDGREVIFYKIIDYILHGKEIKVPTPPMDHWTLSISGLPTTYAENG 352
QY 50 DPPELLSAPQDVKGHCHEAFAFQAKLKLP-----SNPKNKTFIDLVLAQLRRRL 100
DB 353 ICNMLNAPAD-----EYFTFKQGPVDEGTGWIKNVLSPVKNKEDIGVATFYNRK 404
QY 101 PARRGGKKQKHIK-----CPSCDSYEKRTPKFEFLERLKWLLQKHIH 144
DB 405 DGKPPDEYDEHIAETLTQFLGWSLLNTDTYKMNK----LENRKDIAQEMLMNH 454

RESULT 12
```

```

IL15_BOVIN
ID IL15_BOVIN STANDARD; PRT; 162 AA.
AC Q28028;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.

OS Bos taurus (Bovine).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HOLSTEIN;
RX MEDLINE=97426124; PubMed=9282828;
RA Canals A., Gasbarre L.C., Boyd P.C., Almeria S., Zarlenga D.S.;
RT "Cloning and expression of bovine interleukin-15: analysis and
RT modulation of transcription by exogenous stimulation.";
RL J. Interferon Cytokine Res. 17:473-480(1997)
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U42433; AAA85130.1; .
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48 POTENTIAL.
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 127 127 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 162 AA; 18554 MW; 6A63CAA329EB8302 CRC64;

Query Match 8.9%; Score 68.5; DB 1; Length 162;
Best Local Similarity 21.8%; Pred. No. 4.9;
Matches 24; Conservative 21; Mismatches 58; Indels 7; Gaps 3;

QY 38 IVPLKTIYEN-----DLDPKSAPODVKGCHFAAFACFOKAKLPSPGNKNTFTIDL 92
Db 53 VINDLKTIEHLIOSIHMDATLYT-ESDAHPNCKVTAMQCFLELRVILHESKNAT-IYEI 110

QY 93 VAOLRRRLPARRGGKKOKHIAKPCSDSYEKRTPKFELERLKLWLLQKMIH 142
Db 111 IENLTWLANSLSSNIENKTELGLCKECEELEEKSEIKSEKFLKSFVHVQMFIN 160

RESULT 13
IL15_CERAE
ID IL15_CERAE STANDARD; PRT; 162 AA.
AC P40221;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Cercopithecus aethiops (Green monkey) (Grivet).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 49-81.
RC TISSUE=Kidney;
RX MEDLINE=94233380; PubMed=8178155;
RA Grabstein K.H., Eisenman J., Shanebeck K., Rauch C.,
RA Srinivasan S., Fung V., Beers C., Richardson J., Schoenborn M.A.,
RA Ardien M., Johnson L., Alderson M.R., Watson J.D., Anderson D.M.,
RA Giri J.G.;
RT "Cloning of a T cell growth factor that interacts with the beta chain
RT of the interleukin-2 receptor.";
RL Science 264:965-968(1994).
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U03099; AAA18416.1; .
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 48 POTENTIAL.
FT CHAIN 49 162 INTERLEUKIN-15.
FT DISULFID 83 133 POTENTIAL.
FT CARBOHYD 90 136 POTENTIAL.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 162 AA; 18222 MW; 1BF9A82644E1C9B7 CRC64;

Query Match 8.9%; Score 68.5; DB 1; Length 162;
Best Local Similarity 22.8%; Pred. No. 4.9;
Matches 26; Conservative 22; Mismatches 53; Indels 13; Gaps 4;

QY 32 LRHLIDIVEQKIYENDLDPKLSAPQDVKGCHFAAFACFOKAKLPSPGNKNTFTID 91
Db 57 LKKIEDLIQSMHI-----DATLYTESDVHPSPKVTAMKCF-LLEQLVISHSGDTIDH 109

QY 92 LVAQ---LRRRLPARRGGKKOKHIAKPCSDSYEKRTPKFELERLKLWLLQKMIH 142
Db 110 TVENLIILANLSSNGNITE---SGCKECEELEEKNIKEFLQSPFVHVQMFIN 160

RESULT 14
IL15_MACMU
ID IL15_MACMU STANDARD; PRT; 162 AA.
AC P48092;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE INTERLEUKIN-15 PRECURSOR (IL-15).
GN IL15.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=96003435; PubMed=7561102;
RA Villingner F.J., Brar S.S., Mayne A.E., Chikkala N., Ansari A.A.;

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 23, 2001, 11:42:48 ; Search time 96.38 Seconds
(without alignments)
177.551 Million cell updates/sec

Title: US-09-522-217-56

Perfect score: 769

Sequence: 1 MERTLVCLVIFLGTVAHKS.....KEFLRLKWLQKMIHQHLS 146

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL15.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_unclassified.*

13: sp_vertebrate.*

14: sp_virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85.5	11.1	476	13	Q91940 xiphophorus
2	84	10.9	203	5	Q9N71 leishmania
3	83.5	10.9	513	5	Q9N89 trypanosoma
4	82.5	10.7	414	14	Q6661 equine herp
5	82	10.7	673	10	Q9SA08 arabidopsis
6	79.5	10.3	1304	10	Q9SUK4 arabidopsis
7	78.5	10.2	404	5	Q9XWF2 caenorhabdi
8	78.5	10.2	440	10	Q9M2E9 arabidopsis
9	77	10.0	143	13	Q9IAC7 g9iac7 meleagris g
10	76	9.9	566	2	Q86854 streptomyce
11	76	9.9	797	2	Q9RKS9 streptomyce
12	75.5	9.8	372	13	Q93327 gallus gall
13	75.5	9.8	494	5	Q9VWU0 drosophila
14	75	9.8	718	1	O27661 methanobact
15	75	9.8	1054	5	Q9W493 arabidopsis
16	75	9.8	1361	11	O04264 arabidopsis
17	74.5	9.7	135	11	O88648 marmota mon
18	74	9.6	354	14	O65168 african swi
19	74	9.6	1077	13	Q91019 gallus gall

20	74	9.6	1081	3	O42696 candida alb
21	74	9.6	1081	3	O74271 candida alb
22	74	9.6	1081	3	O93851 candida alb
23	74	9.6	1081	3	Q9URL9 candida alb
24	74	9.6	1142	13	Q02391 gallus gall
25	73.5	9.6	818	4	Q9UP55 homo sapien
26	73	9.5	474	5	Q9N9T8 leishmania
27	72.5	9.4	372	11	Q9QZQ8 mus musculu
28	72.5	9.4	698	4	Q9Y4J6 homo sapien
29	72.5	9.4	771	4	Q9Y462 homo sapien
30	72.5	9.4	835	14	P88992 murid herpe
31	72.5	9.4	1891	4	Q9NNY1 homo sapien
32	72.5	9.4	1912	4	Q9ULL4 homo sapien
33	72	9.4	657	10	Q9ZUT5 arabidopsis
34	72	9.4	991	5	Q9VIF5 drosophila
35	71.5	9.3	403	1	O26176 methanobact
36	71.5	9.3	487	6	Q9N2B2 pan troglod
37	71.5	9.3	487	6	Q9N2B1 gorilla gor
38	71	9.2	409	2	Q9K9A8 bacillus ha
39	71	9.2	488	2	O87810 pseudomonas
40	71	9.2	2420	10	Q9SL90 arabidopsis
41	70.5	9.2	223	13	Q93325 gallus gall
42	70.5	9.2	253	2	O66717 aquifex aeo
43	70.5	9.2	324	5	Q9Y028 trypanosoma
44	70.5	9.2	369	13	Q93326 gallus gall
45	70.5	9.2	372	4	O75367 homo sapien

ALIGNMENTS

RESULT 1

ID Q91940 PRELIMINARY; PRT; 476 AA.

AC Q91940;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE 163A ZFI (ZFI).

GN ZFI.

OS Xiphophorus maculatus (Southern platyfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;

OC Cyprinodontiformes; Poeciliidae; Xiphophorus.

OX NCBI_TaxID=8083;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=163A; TISSUE=MUSCLE;

RX MEDLINE=96113537; PubMed=85950505;

Query Match 11.1%; Score 85.5; DB 13; Length 476;
Best Local Similarity 22.0%; Pred. No. 1.1;
Matches 37; Conservative 20; Mismatches 42; Indels 69; Gaps 8;

QY 18 HKSSPQGPQRLLR-----LRHLIDI-----VFQKLYE 46
|:|:| |:|:|
Db 171 HESSPLGSKLIVKREPKLHCKYCYETAEOGLLNRLHLAVHSKNFAHVCVCAKFR 230
|:|:| |:|:|
QY 47 NLDPELLSAPQDVKGH-----CEHAAPACQKAKLK---PSNPGNNKTFIIDL 92

```

Db 231 H-----PSELKHMRTHTGKPYQCPCFCRCADQSNLTKTHKSKHGADLPFCRNH 281
QY 93 VAOLRRRLPARRGGKKQKHIA-----KPCSDSYEKRTPKFEFLER 132
Db 282 CPQ-----PTPDARELQRIHWQGHKTHOCPHCE--HKSTNSSDLAKR 322

RESULT 2
Q9NN71 PRELIMINARY; PRT; 203 AA.
AC Q9NN71;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE PROBABLE AXONEME CENTRAL APPARATUS PROTEIN (FRAGMENT).
GN LM15.350.
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL160371; CAC00339.1; -.
FT NON_TER 1
FT NON_TER 203
SQ SEQUENCE 203 AA; 22391 MW; 29E3EFD8BF87EC9A CRC64;

Query Match 10.9%; Score 84; DB 5; Length 203;
Best Local Similarity 23.4%; Pred. No. 0.64;
Matches 32; Conservative 26; Mismatches 59; Indels 20; Gaps 4;

QY 5 LVCLVIFLGTVAHKSSPQGDRLRLRLHLDIVLQKLIYENDLDPQL-LSAPQDVKGH 63
Db 74 LLRLLDVYL-----SPSSDDLDQTKSKRAKAIHQCVLPALEPLHPADQDV--- 123
QY 64 CEHAAFACFOKALKPSNPNKNTFIIDLVAQLRRRLPARRGGKKQKHIAKPCSC----- 118
Db 124 ---LKVYCGQYAKVLDVAAREFVANRGLATVQRIKAEPGSSLAESIQLIINSCFPPEI 180

QY 119 -DSYKERTPKFEFLERLK 134
Db 181 VEYSPYAQTFFIEKIE 197

RESULT 3
Q9NN89 PRELIMINARY; PRT; 513 AA.
AC Q9NN89;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE PROBABLE AXONEME CENTRAL APPARATUS PROTEIN.
GN CHRI.235.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TRE0927;
RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,
RA Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,
RA Gerrard C., Rajandream M.A., Barrell B.G.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL359782; CAB95504.1; -.
SQ SEQUENCE 513 AA; 56126 MW; 40BB4546A1B390C3 CRC64;

Query Match 10.9%; Score 83.5; DB 5; Length 513;
Best Local Similarity 23.3%; Pred. No. 1.9;
Matches 31; Conservative 26; Mismatches 57; Indels 19; Gaps 4;

QY 22 PQGPDRLRLIR-----LRHLIDIVEQLKIYENDLDPQLLSAPQDVKGHCEHAFAFCQKAKL 77
Db 390 PSSDDLEMKSKRAKNIHQCVLPALPLHPE---APKNV-----LKVYCGQFAKV 440
QY 78 KPSNPNKNTFIIDLVAQLRRRLPARRGGKKQKHIAKPCSC-----DSYKERTPKFEFLER 131
Db 441 LPTDIAAKREFVANRGLATVQRIQPEPGSKLAEYIQSINNCCYPPEIVQYVSPQYATPFL 500

RESULT 4
Q66661 PRELIMINARY; PRT; 414 AA.
AC Q66661;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE DNA POLYMERASE PROCESSIVITY SUBUNIT.
OS Equine herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Roseolovirus.
OX NCBI_TaxID=12657;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=86/67;
RX MEDLINE=95302501; PubMed=7783207;
RA Telford E.A., Watson M.S., Aird H.C., Perry J., Davison A.J.;
RT "The DNA sequence of equine herpesvirus 2.";
RL J. Mol. Biol. 249:520-528(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=86/67;
RA Telford E.A.R.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20824; AAC13847.1; -.
SQ SEQUENCE 414 AA; 45191 MW; BA5A691785BE441F CRC64;

Query Match 10.7%; Score 82.5; DB 14; Length 414;
Best Local Similarity 25.5%; Pred. No. 1.9;
Matches 28; Conservative 16; Mismatches 51; Indels 15; Gaps 3;

QY 19 KSSPQGDRLRLRLHLDIVLQKLIYENDLDPQLLSAPQDVKGHCEHAFAFCQKAKLK 78
Db 281 KSSKPASGELSVELLRANPQVDFNGVPEGDVQTEVSS---VASTCRHLSESCSLDPPT 337
QY 79 PSNPNKNTFIIDLVAQLRRRLPARRGGKKQKHIAKPCSDSYKERTPK 128
Db 338 PELFGSPDTF-----KEIPGRSG---SVHLERDLSCSDSEETPKQ 375

RESULT 5
Q9SA08 PRELIMINARY; PRT; 673 AA.
AC Q9SA08;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE F28K20.8 PROTEIN.
GN F28K20.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
```


RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
RA Li J., Krenetskaia I., Luros J., Gonzalez A., Alkafi H., Araujo R.,
RA Chao Q., Connell A.B., Dunn P., Hansen N., Huizar L., Kim C.,
RA Palm C.J., Rowley D., Shinn P., Walker M., Davis R.W., Ecker J.R.,
RA Federspiel N.A., Theologis A.;
RT "Arabidopsis thaliana chromosome 1 BAC F28K20 sequence.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN [2]
RC
RC SEQUENCE FROM N.A.
RC STRAIN=CV. COLOMBIA;
RL Theologis A.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN [3]
RC
RC SEQUENCE FROM N.A.
RC STRAIN=CV. COLOMBIA;
RL Theologis;
RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AC004793; AAC21698.1; -
SQ SEQUENCE 673 AA; 76895 MW; 8718AA7CF6E002CD CRC64;

Query Match	10.7%	Score 82;	DB 10;	Length 673;
Best Local Similarity	23.8%;			
Pred. No. 3.7;				
Matches 39;	Conservative	28;	Mismatches 43;	Indels 54;
Gaps				9;

[illegible]

RESULT	6	
Q95UK4		
ID	Q95UK4	PRELIMINARY; PRT; 1304 AA.
AC	Q95UK4;	
DT	01-MAY-2000 (TREMBLrel. 13, Created)	
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)	
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)	
DE	DISEASE RESISTANCE RPP5 LIKE PROTEIN.	
GN	DL4490C OR AT4G16920.	
OS	Arabidopsis thaliana (Mouse-ear cress).	
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophytes	
OC	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; e	
OC	Brassicales; Brassicaceae; Arabidopsis.	
OX	NCBI_TaxID=3702;	
OX	[1]	
RP	SEQUENCE FROM N.A.	
RP	Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., T.	
RA	Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,	
RA	Puigdomenech P., Hatzopoulos P., Obermaier B., Duesterho	
RA	Jones J., Palme K., Ansgorge W., Delseny M., Bancroft I.,	
RA	Schuelter C., Chalwatzis N.	
RA	Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RP	EU Arabidopsis sequencing project;	
RN	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RP	EU Arabidopsis sequencing project;	
RA	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; Z57342; CAB46046.1; -; -;	
DR	EMBL; AL161545; CAB80963.1; -;	
DR	INTERPRO; IPR000157; -;	
DR	INTERPRO; IPR000767; -;	
DR	INTERPRO; IPR001611; -;	

```

DR INTERPRO: IPR002182; -.
DR PFAM: PF00560; LRR; 4.
DR PFAM: PF00931; NB-ARC; 1.
DR PFAM: PF01582; TIR; 1.
DR PRINTS: PR00364; DISEASERSIST.
SQ SEQUENCE 1304 AA; 147728 MW; BE661972EF2DD0DA CRC64;

      Query Match      10.3%; Score 79.5; DB 10; Length 1304;
      Best Local Similarity 23.7%; Pred. No. 13;
      Matches 41; Conservative 28; Mismatches 53; Indels 51; Gaps

Qy 12 FLGVVAKHSSPQG-----PDLRLIRLHUIDIVQOLKIYENDLDPELL----- 54
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 297 FLKTLVGKAEWFGSGRIIVITQDROLKLA-AHEIDLVEYEVKLPQSGLALQLMISQYAFGKD 355
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

Qy 55 SAPQDVKGHCHEHAFAFCQAKLKPSNP-----GNKKTFFIDLVAQLRR----- 98
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 356 SPDDDFKA-----LAFEVAGLAGSLPLGLSVGLSGLRGDKDEWVKWMPRLRNDSDDK 408
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

Qy 99 -----RUPARRGGKKOKHIAKPCSC--DSYEKRTPKFELE--RLKWLQLQKMI 141
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 409 IEETLRVCYDLRNLKKNRELKFCIACFNGFKVSNVRELLEDDVGLTMLVEKSL 461
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:

RESULT 7
Q9XWF2 PRELIMINARY; PRT; 404 AA.
AC Q9XWF2;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)

```

RESULT	7	
Q9XWF2		
ID	Q9XWF2	PRELIMINARY; PRT; 404 AA.
AC	Q9XWF2;	
DT	01-NOV-1999	(TrEMBLrel. 12, Created)
DT	01-NOV-1999	(TrEMBLrel. 12, Last sequence update)
DT	01-JUN-2000	(TrEMBLrel. 14, Last annotation update)
DE	Y52B11a.9	PROTEIN.
DE	Y52B11a.9	
GN	Caenorhabditis elegans.	
OS	Caenorhabditis elegans.	
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea	
OC	Rhabditidae; Peloderinae; Caenorhabditis.	
OX	NCBI_TaxID=6239;	
OX	[1]	
RN	SEQUENCE FROM N.A.	
RP	MEDLINE=94150718; PubMed=7906398;	
EX	Lennard N.;	
RA	"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.	
RT	elegans.";	
RT	Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.	
RL		

RA MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
 RA Smalton N., Smith A., Sonnhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.; et al.
 RA *2.2 Mb of contiguous nucleotide sequence from chromosome III of *C.*
 NT *elegans*.
 NT Nature 368:32-38(1994).
 RL ENBL; AL032654; CAA21720.1; -.
 DR INTERPRO: IPR000822; -.
 DR INTERPRO: IPR002358; -.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; UNKNOWN_1.
 DR PROSITE: PS00525; RIBOSOMAL_L6_1; 1.
 SQ PROSEQUENCE 404 AA; 46400 MW; 9F55A1413525F057 CRC64;

Query Match	10.2%	Score 78.5;	DB 5;	Length 404;
Best Local Similarity	23.8%	Pred. No. 4.8;		
Matches	34;	Conservative	20;	Mismatches 58;
				Indels 31;
				Gaps 6;

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OX Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
NCBI_TaxID=1902;
[1]
SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Oliver K., Harris D.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cardeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
Kinashi H., Hopwood D.A.;
RT A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomycetes coelicolor A3(2) chromosome.;
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL132824; CAB60182.1; -;
KW Hypothetical protein.
SQ SEQUENCE 797 AA; 81895 MW; 8A13FD0843726EAD CRC64;

Query Match 9.9%; Score 76; DB 2; Length 797;
Best Local Similarity 27.7%; Pred. No. 18;
Matches 36; Conservative 13; Mismatches 39; Indels 42; Gaps 6;

QY 19 KSPQGGPQRLRLRLHLD:-----IVEQLKIYENDLDPE-----LLS 55
DB 451. RGSRSRYQVRLTLGSDWDRLDAAVERPCHIAALLDGLPHSLADLRGVPLP 510
QY 56 APQDV-----KGH-CEHAACFQKAKLKPSNP-----GNKTFIDLVQALRRR 99
DB 511 GPGDLAPRCSPDSGHPCKHAALCYOTARLLDADPEVLLLRGRGERALLD---ALSRR 567
QY 100 LPARRGGKKQ 109
DB 568 NAAREARAQ 577

RESULT 12
ID Q93327 PRELIMINARY; PRT; 372 AA.
AC Q93327;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE HISTONE MACROH2A.2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA MEDLINE=98278895; PubMed=9611225;
RX Pehrson J.R., Fuji R.N.;
RT "Evolutionary conservation of histone macroH2A subtypes and domains.";
RL Nucleic Acids Res. 26:2837-2842(1998).
DR EMBL; AF058446; AAC28847.1; -;
DR INTERPRO; IPR000166; -;
DR INTERPRO; IPR002119; -;
DR INTERPRO; IPR002589; -;
DR PFAM; PF00125; histone; 1.
DR PFAM; PF01661; DUF27.1.
DR PRINTS; PR00620; HISTONEH2A.
DR PRODOM; PD000565; -; 1.
SQ SEQUENCE 372 AA; 39656 MW; 8F5940BCD57C6E95 CRC64;

QY 48 DLDPELLAPQDVKGCHHAAPACFQKAKLKPSNPGNKKTFIDLVQALRRRLPARRGK 107
DB 108 NIHPPELLAKRSGKLE--AIITPPAK-KAKSPQKKT----VSKTGGKGARKSKK 160
QY 108 KOKHIACPCSDSYEKTRPKPELRK----WLLQK--MIHQHLS 146
DB 161 KQGEVSKSASADSTTEGTPADGFTVLSTKSLFGORLNLIHSEIS 205

RESULT 13
ID Q9VWU0 PRELIMINARY; PRT; 494 AA.
AC Q9VWU0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CG15046 PROTEIN.
GN CG15046.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.G., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan K., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003509; AAP48847.1; -;
DR FLYBASE; FBgn0030927; CG15046.
DR INTERPRO; IPR001254; -;

QY 49 LDPILLSAPOVKCHCEHAFACFOKAKLXPSNPGNNKTFIIDLVAOLRRRLPARRGGKK 108
Db 188 IEPSEHEHPHVNGH-----PHVVFORSSVVKPH-----SLRKNRKRKGGS 230
QY 109 QKHIACPCSDSYEKRTPKFEFLERLKW 135
Db 231 GSG-AEVSNCGTREPRRME--TRLEW 254

Search completed: May 23, 2001, 11:49:00
Job time: 372 sec